

# NUMERICAL SIMULATION OF POLYNOMIAL-SPEED CONVERGENCE PHENOMENON

YAO LI AND HUI XU

**ABSTRACT.** We provide a hybrid method that captures the polynomial speed of convergence and polynomial speed of mixing for Markov processes. The hybrid method that we introduce is based on the coupling technique and renewal theory. We propose to replace some estimates in classical results about the ergodicity of Markov processes by numerical simulations when the corresponding analytical proof is difficult. After that, all remaining conclusions can be derived from rigorous analysis. Then we apply our results to two 1D microscopic heat conduction models. The mixing rate of these two models are expected to be polynomial but very difficult to prove. In both examples, our numerical results match the expected polynomial mixing rate well.

## 1. INTRODUCTION

The aim of this paper is two-fold. From the viewpoint of statistical mechanics, this paper aims to justify the polynomial ergodicity of a class of 1D microscopic heat conduction models. Purely rigorous analysis of polynomial ergodicity of these models using current analytical techniques fails to provide satisfactorily accurate results. This paper also aims to establish a comprehensive method that can be applied on a broader scale. That is, from the viewpoint of numerical analysis, we also want to propose a hybrid method that captures polynomial-speed convergence to steady-states for general Markov processes.

Heat conduction is ubiquitous in the universe and has been well-studied at the macroscopic level. However, from a microscopic point of view, the study of how energy is transported in materials is a very challenging topic. In particular, the derivation of Fourier's law from microscopic Hamiltonian dynamics is a century-old challenge to mathematicians and physicists [1, 7, 8, 9, 24, 2]. Due to the significant difficulty of analyzing Hamiltonian models, many researchers seek stochastic approximations of Hamiltonian dynamics in microscopic heat conduction models [8, 17, 9, 14, 10, 6]. In this paper, we will work primarily on these stochastic heat conduction models.

When an 1D stochastic heat conduction model is connected to two thermalized boundaries with different temperatures, one would expect the existence of a naturally selected steady-state, called the non-equilibrium steady-state (NESS). An analysis of topics like existence and uniqueness of NESS, and speed of convergence to NESS will open the door to further studies such as the thermal conductivity, the existence of local thermodynamic equilibrium, the fluctuation theorem and Gallavotti-Cohen

---

Hui Xu was supported in part by REU program of University of Massachusetts Amherst.

symmetry, and eventually the Fourier's law. However, a rigorous analysis about the ergodicity of the NESS is usually very challenging.

Due to complicated interactions within the chain, a stochastic microscopic heat conduction model may have sub-exponential speed of mixing and sub-exponential speed of convergence to the NESS. In this paper we will present two 1D microscopic heat conduction models, namely, the stochastic energy exchange model and the random halves model, both of which originate from deterministic dynamical systems [8, 3, 17]. One common feature of these two models is that a low energy particle (or a low energy site) requires a long time to have the next energy exchange. Because of this, we expect the rate of convergence and the rate of mixing to be  $\sim t^{-2}$ . We refer readers to Section 5 and Section 6 for more engaged discussion about microscopic heat conduction models and their connections to deterministic dynamical systems. Besides these two models, other microscopic heat conduction models that have sub-exponential mixing rate include the particle model in [27, 26], the rotor model in [4, 5], and the anharmonic chain in [12]. Other examples of sub-exponential rate of convergence have also been seen in various models like MCMC algorithms and random walks [13, 25, 20].

Regardless of the detailed setting of models, a rigorous proof of slow mixing phenomena is known to be very difficult, partially because all methods that use the spectral gap of the infinitesimal generator simply fail to work for these Markov processes. Even if a rigorous proof is possible in simpler models such as the stochastic energy exchange model studied in this paper [16], the bound of convergence speed to the steady-state is usually not accurate, because an explicit expression of the NESS is usually not possible.

A numerical justification of slow mixing (and slow convergence) phenomenon is challenging too. Let  $P^t$  be the transition kernel of a Markov process  $X_t$ . Let  $f$  and  $g$  be two observables on the state space of  $X_t$ . The decay of correlation is denoted by

$$C_\mu(t) = \left| \int (P^t f)(x)g(x)\mu(dx) - \int (P^t f(x)\mu(dx) \int g(x)\mu(dx) \right|,$$

where  $\mu$  is a probability measure. A direct simulation of  $C_\mu(t)$  requires Monte Carlo simulations of  $\int (P^t f)(x)g(x)\mu(dx)$  and  $\int (P^t f(x)\mu(dx)$ . Therefore, it is easy to see that the estimator of  $C_\mu(t)$  has variance  $O(1)$ . Now assume  $C_\mu(t)$  has a polynomial tail  $C_\mu(t) \sim t^{-\alpha}$ . A simple calculation shows that to make the relative error of  $C_\mu(T)$  be less than  $\epsilon$ , the sample size should be at least  $\epsilon^{-2}T^{2\alpha}$ , which brings the total computational cost to  $\epsilon^{-2}T^{2\alpha+1}$ . Our simulation shows that large  $T$  is usually necessary to effectively capture the tail of  $C_\mu(t)$ , which makes the computational cost of direct Monte Carlo simulation not acceptable. For instance, when  $T = 10^3$ ,  $\alpha = 2$ , and  $\epsilon = 0.1$ , the computational cost of direct simulation is  $\sim 10^{17}$ .

In this paper, we present a hybrid method that combines the advantage of both analytical and numerical methods to calculate the polynomial speed of convergence for Markov processes. As an application of this method, we numerically show that the two microscopic heat conduction models have speed of convergence  $\sim t^{-2}$  to

their steady states. This is consistent to both our heuristic analysis and numerical results for corresponding deterministic models.

When spectral methods fail to work, a general approach to prove speed of convergence to steady-states is based on the coupling method and discrete renewal theory. We run two independent copies of Markov processes until they are coupled and become indistinguishable. The rule of coupling is that when both processes enter a certain set  $\mathfrak{C}$  called the uniform reference set (or small set in some literatures), they have positive probability to couple. Then the coupling lemma tells us that the speed of convergence in the total variation norm is mainly determined by the tail of the first passage time to  $\mathfrak{C}$ , denoted by  $\tau_{\mathfrak{C}}$ . It is usually not difficult to construct such a uniform reference set  $\mathfrak{C}$ . But an analytical estimation of  $\tau_{\mathfrak{C}}$  is usually difficult. The main strategy proposed in this paper is to numerically estimate  $\tau_{\mathfrak{C}}$ . Different from the correlation function, the first passage time of a Markov chain can be computed easily with very high accuracy. After obtaining a numerical tail of  $\tau_{\mathfrak{C}}$ , we use coupling technique and renewal theory to show that the Markov process has a polynomial speed of convergence.

The organization of this paper is as follows. Section 2 serves as a probability preliminary. Then we will discuss numerically or analytically verifiable conditions that lead to a polynomial convergence rate in Section 3. Section 4 discusses conclusions that can be made from numerical and analytical conditions. Lastly, we study two microscopic heat conduction models in Section 5 and Section 6, respectively.

## 2. PROBABILITY PRELIMINARY: CONVERGENCE RATES OF MARKOV CHAINS

The purpose of this section is to review known sufficient conditions towards the polynomial ergodicity of a Markov process. Throughout this section, we let  $\Psi_n$  be a discrete-time Markov chain on a measurable space  $(X, \mathcal{B})$ . The transition kernel of  $\Psi_n$  is denoted by  $\mathcal{P}(x, \cdot)$ . For each  $A \in \mathcal{B}$ ,  $P(\cdot, A)$  is a measurable function. For each  $x \in X$ ,  $P(x, \cdot)$  is a probability measure.

For  $A \in \mathcal{B}$ , we let  $\tau_A$  be the first passage time to  $A$ :

$$\tau_A = \inf\{n > 0 \mid \Psi_n \in A\}.$$

A set  $A \in \mathcal{B}$  is said to be accessible if  $\mathbb{P}_x[\tau_A < \infty] = 1$  for every  $x \in X$ .

We say a Markov process is *irreducible* with respect to a measure  $\phi$  on  $\mathcal{B}$  if every  $A \in \mathcal{B}$  with  $\phi(A) > 0$  is accessible. We refer readers to Chapter 4 of [21] for detailed definition and properties of irreducibility. If  $\Psi_n$  is irreducible with respect to a non-trivial measure  $\phi$ , then there exists a “maximal irreducible measure”  $\psi$  such that  $\phi \ll \psi$ , where  $\psi$  is unique up to equivalence classes. We skip the formal introduction of the maximal irreducibility as the  $\phi$ -irreducibility is sufficient for this paper.

**(A) Construction of an atom.** It has long been known that the stochastic stability of  $\Psi_n$ , such as recurrence, ergodicity, and decay of correlation, follows from certain pseudo-atomic properties [21, 23]. More precisely, we need a *uniform reference set*  $\mathfrak{C} \in \mathcal{B}$  that satisfies

$$\sup_{x \in \mathfrak{C}} \mathcal{P}(x, \cdot) \geq \eta \theta(\cdot),$$

where  $\theta$  is a probability measure on  $(X, \mathcal{B})$  and  $\eta$  is a strictly positive real number.

Assuming the existence of such a uniform reference set  $\mathfrak{C}$ , the state space of  $\Psi_n$  can be splitted by letting  $\tilde{X} = X \cup \mathfrak{C}_1$ , where  $\mathfrak{C}_1$  is an identical copy of  $\mathfrak{C}_0 := \mathfrak{C}$ . Then we can naturally extend  $(X, \mathcal{B})$  to  $(\tilde{X}, \tilde{\mathcal{B}})$  and “split” a probability measure  $\mu$  into a probability measure  $\mu^*$  on  $(\tilde{X}, \tilde{\mathcal{B}})$ :

$$\begin{cases} \mu^*|_X = (1 - \eta) \mu|_{\mathfrak{C}_0} + \mu|_{X \setminus \mathfrak{C}_0} \\ \mu^*|_{\mathfrak{C}_1} = \eta \mu|_{\mathfrak{C}_0}, \quad \mathfrak{C}_0 \cong \mathfrak{C}_1 \text{ via the natural identification.} \end{cases}$$

With the above construction, we can define the split chain  $\tilde{\Psi}$  on  $(\tilde{X}, \tilde{\mathcal{B}})$  with a transition kernel  $\tilde{\mathcal{P}}(x, \cdot)$ :

$$\begin{cases} \tilde{\mathcal{P}}(x, \cdot) = (\mathcal{P}(x, \cdot))^* & x \in X \setminus \mathfrak{C}_0 \\ \tilde{\mathcal{P}}(\mathbf{x}, \cdot) = [(\mathcal{P}(x, \cdot))^* - \eta\theta^*(\cdot)]/(1 - \eta) & x \in \mathfrak{C}_0 \\ \tilde{\mathcal{P}}(x, \cdot) = \theta^*(\cdot) & x \in \mathfrak{C}_1 \end{cases}$$

It is straightforward to check that  $\tilde{\Psi}_n$  possesses an atom  $C_1$ : for all  $x \in C_1$ ,  $\mathcal{P}(x, \cdot) = \theta(\cdot)$ . In addition, the natural projection  $\pi : \tilde{X} \rightarrow X$  projects  $\tilde{\Psi}_n$  to  $\Psi_n$ . This split construction is called the Nummelin splitting. We refer [22] for the detail.

The aperiodicity of  $\Psi_n$  can follow from the properties of  $\mathfrak{C}$ . If  $\Psi_n$  is irreducible and admits a uniform reference set such that  $\theta(\mathfrak{C}) > 0$ ,  $\Psi_n$  is said to be *strongly aperiodic*. If  $\Psi_n$  is strongly aperiodic,  $\Psi_n$  must be aperiodic such that no cyclic decomposition is possible. We refer readers to Chapter 5 of [21] for the complete statement of aperiodicity of Markov processes.

The following theorem gives the existence of an invariant probability measure of  $\Psi_n$ .

**Theorem 2.1** (Theorem 10.0.1 from [21]). *Let  $\Psi_n$  be an irreducible aperiodic Markov process on  $(X, \mathcal{B})$ . If  $\mathfrak{C} \in \mathcal{B}$  is an accessible uniform reference set such that*

$$\sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}] < \infty$$

*then there exists an invariant probability measure  $\pi$ .*

**(B) Coupling.** The speed of convergence of  $\Psi_n$  then follows from the following coupling argument. Without loss of generality, assume that the uniform reference set  $\mathfrak{C}$  is accessible, i.e.,

$$\mathbb{P}_x[\tau_{\mathfrak{C}} < \infty] := \mathbb{P}_x[\inf\{n > 0 \mid \Psi_n \in \mathfrak{C}\}] = 1$$

for any  $x \in X$ . Let  $\mu$  and  $\nu$  be two initial distributions. One way to bound  $\|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{TV}$  as  $n \rightarrow \infty$  is to run two independent copies of  $\tilde{\Psi}_n$  starting from  $\mu^*$  and  $\nu^*$ , respectively, and perform a coupling at their first simultaneous return to the atom  $\mathfrak{C}_1$ . Let  $T$  be the coupling time. It is well known that

$$\|\mu\mathcal{P}^n - \nu\mathcal{P}^n\|_{TV} \leq \|\mu^*\tilde{\mathcal{P}}^n - \nu^*\tilde{\mathcal{P}}^n\|_{TV} \leq 2\mathbb{P}[T > n].$$

Let  $Y_0, Y_1, Y_2, \dots$  and  $Y'_0, Y'_1, Y'_2, \dots$  be the passage times to  $\mathfrak{C}_1$  for two independent processes, respectively. It is obvious that  $Y_1, Y_2, \dots$  and  $Y'_1, Y'_2, \dots$  are *i.i.d* random

variables with a distribution  $\tau_{\mathfrak{C}_1}|\mathfrak{C}_1$ . Therefore, the coupling time  $T$  is the first simultaneous renewal time for renewal processes

$$S_n := \sum_{i=0}^n Y_i, \quad \text{and} \quad S'_n := \sum_{i=0}^n Y'_i.$$

More precisely, we have

$$T = \inf_{n \geq 0} \{S_i = S'_j = n \text{ for some } i, j\}.$$

If in addition, the return times to  $\mathfrak{C}_1$  are aperiodic, i.e.,  $\gcd\{n \geq 1 \mid \mathbb{P}[Y_i = n] > 0\} = 1$ , then it follows from [19] that the finiteness of the moments of  $T$  is implied by the finiteness of corresponding moments for  $Y_0, Y'_0$ , and  $Y_1$ .

**Theorem 2.2** (Theorem 4.2 from [19]). *Let  $S_n$  and  $S'_n$  be the renewal processes as above. If there exists  $\beta > 1$  such that  $\mathbb{E}[Y_0^\beta]$ ,  $\mathbb{E}[Y'_0{}^\beta]$ , and  $\mathbb{E}[Y_1^\beta]$  are all finite, then there exists a constant  $C$  such that*

$$\mathbb{E}[T^\beta] \leq C(\mathbb{E}[Y_0^\beta] + \mathbb{E}[Y'_0{}^\beta]) < \infty.$$

Theorem 2.2 implies the following immediately:

**Theorem 2.3.** *Let  $\Psi_n$  be a Markov chain on  $(X, \mathcal{B})$  with transition kernel  $\mathcal{P}$ . Suppose  $\Psi_n$  has an atom  $\mathfrak{C}_1$  that is accessible and whose return times are aperiodic. Let  $\mu$  and  $\nu$  be two probability distributions on  $X$ , and assume that for some  $\beta > 1$ ,*

$$\mathbb{E}_\mu[\tau_{\mathfrak{C}_1}^\beta], \quad \mathbb{E}_\nu[\tau_{\mathfrak{C}_1}^\beta] \quad \text{and} \quad \mathbb{E}_{\mathfrak{C}_1}[\tau_{\mathfrak{C}_1}^\beta] < \infty.$$

Then

$$\lim_{n \rightarrow \infty} n^\beta \|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{\text{TV}} = 0.$$

The speed of convergence follows immediately by applying Theorem 2.3 to the split chain  $\tilde{\Psi}_n$ .

### (C) Convergence rate for general Markov chain.

It remains to pass the result of  $\tilde{\Psi}_n$  to  $\Psi_n$ . Note that if Theorem 2.3 holds for  $\tilde{\Psi}_n$ , we have

$$\lim_{n \rightarrow \infty} n^\beta \|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{\text{TV}} \leq \lim_{n \rightarrow \infty} n^\beta \|\mu^* \tilde{\mathcal{P}}^n - \nu^* \tilde{\mathcal{P}}^n\|_{\text{TV}} = 0.$$

Therefore, result for  $\Psi_n$  follows from the following lemma that passes bounds of  $\tau_{\mathfrak{C}}^\beta$  to bounds of  $\tau_{\mathfrak{C}_1}^\beta$ .

**Lemma 2.4** (Lemma 3.1 of [23]). *Let  $\Psi_n$  be an aperiodic Markov chain on  $(X, \mathcal{B})$ . If  $\mathfrak{C} \in \mathcal{B}$  is an accessible uniform reference set and*

$$\sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}^\beta] < \infty$$

*for some  $\beta > 0$ , then for any probability measure  $\mu$  such that  $\mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] < \infty$ , we have  $\mathbb{E}_{\mu^*}[\tau_{\mathfrak{C}_1}^\beta] < C \mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] < \infty$  for some constant  $C$ .*

In summary, the rigorous result for polynomial rate of convergence is as follows.

**Theorem 2.5** (Theorem 2.7 of [23]). *Let  $\Psi_n$  be an aperiodic Markov chain on  $(X, \mathcal{B})$  with transition kernel  $\mathcal{P}$ . Assume  $\Psi_n$  admits an accessible uniform reference set  $\mathfrak{C}$  such that*

$$\sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}^\beta] < \infty$$

*for some  $\beta > 0$ , then for any probability measures  $\mu, \nu$  on  $X$  that satisfy*

$$\mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] < \infty, \quad \text{and} \quad \mathbb{E}_\nu[\tau_{\mathfrak{C}}^\beta] < \infty,$$

*we have*

$$\lim_{n \rightarrow \infty} n^\beta \|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{TV} = 0.$$

Using more precise bounds in Theorem 2.2 and Lemma 2.4 that involve the initial distributions, we can slightly improve the estimate in Theorem 2.5 in order to estimate the rate of correlation decay.

**Corollary 2.6.** *Let  $\Psi_n$  and  $\mathfrak{C}$  be as in Theorem 2.5. Then for any probability measures  $\mu, \nu$  on  $X$  that satisfy*

$$\mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] < \infty, \quad \text{and} \quad \mathbb{E}_\nu[\tau_{\mathfrak{C}}^\beta] < \infty,$$

*there exists a constant  $C$  such that*

$$\sup_n n^\beta \|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{TV} \leq C(\mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] + \mathbb{E}_\nu[\tau_{\mathfrak{C}}^\beta]).$$

*Proof.* Let  $T$  be the coupling time of  $\tilde{\Psi}_n$ . For any  $n \geq 0$ ,

$$\mathbb{E}_{\mu^*, \nu^*}[T^\beta] \geq \sum_{k=n}^{\infty} k^\beta \mathbb{P}_{\mu^*, \nu^*}[T = k] \geq n^\beta \mathbb{P}_{\mu^*, \nu^*}[T > n].$$

Hence have

$$\sup_n n^\beta \|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{TV} \leq 2 \sup_n n^\beta \mathbb{P}_{\mu^*, \nu^*}[T > n] \leq \mathbb{E}_{\mu^*, \nu^*}[T^\beta].$$

The corollary follows easily from more precise bounds in Theorem 2.2 and Lemma 2.4. □

### 3. VERIFIABLE CONDITIONS FOR SLOW CONVERGENCE

The aim of this section is to convert conditions in Theorem 2.5 to sufficient conditions that are verifiable either numerically or analytically.

Firstly, we will list sufficient conditions that will result in polynomial speed of convergence of  $\|\mu \mathcal{P}^n - \nu \mathcal{P}^n\|_{TV}$  from Theorem 2.5.

- (1)  $\Psi_n$  is irreducible with respect to a non-trivial probability measure  $\phi$ .
- (2)  $\Psi_n$  is aperiodic.
- (3)  $\Psi_n$  admits a uniform reference set  $\mathfrak{C}$  such that

$$\mathcal{P}(x, \cdot) \geq \eta \theta(\cdot) \quad \text{for all } x \in \mathfrak{C}.$$

- (4) There exists a constant  $\beta > 0$  such that  $\mathbb{E}_\mu[\tau_{\mathfrak{C}}^\beta] < \infty$  and  $\mathbb{E}_\nu[\tau_{\mathfrak{C}}^\beta] < \infty$ .
- (5)  $\sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}^\beta] < \infty$  for the constant  $\beta$  in (4).

If in addition, we would like to show the existence of an invariant probability measure  $\pi$  and the polynomial speed of convergence towards  $\pi$ , the following two more conditions are needed.

$$(6) \sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}] < \infty.$$

$$(7) \mathbb{E}_\pi[\tau_{\mathfrak{C}}^\beta] < \infty \text{ for the constant } \beta \text{ in (4).}$$

**(A) Conditions that are verifiable analytically.** For most models, conditions (1)-(3) are relatively easy to check analytically. Condition (1), i.e. the irreducibility, usually can be proved by constructing an event with positive probability such that a positive-measured set is “reachable”. In addition, it is well known that if the Markov process has a continuous component, then the reachability of one point implies the irreducibility.

For many Markov processes, condition (3) is also easy to check. Essentially all we need to show is that the probability measure  $\mathcal{P}(x, \cdot)$  for  $x \in \mathfrak{C}$  has some uniform lower bound, which is usually easy to prove by constructing events with positive probability. It remains to show condition (2), i.e., the aperiodicity. In fact, if  $\mathfrak{C}$  is a uniform reference set with  $\nu(\mathfrak{C}) > 0$ , then  $\Psi_n$  is a strong aperiodic chain, which is obviously aperiodic.

We will give two examples about verifying these analytical conditions in Section 5 and 6.

**(B) Conditions that are easier to check numerically.** It is not difficult to show that a set is a uniform reference set. However, a rigorous estimation of return times to a uniform reference set is usually non-trivial. Most proof relies on the careful construction of a Lyapunov function  $V$ . It is well known that the first passage time to the “bottom” of the Lyapunov function can be estimated by calculating

$$\mathcal{P}V(x) := \int_X \mathcal{P}(x, dy)V(y).$$

Unfortunately, there is no universal approach to construct a Lyapunov function for a Markov process. It may also be nontrivial to prove that a given function is actually a Lyapunov function. We refer [18] for the examples of estimating first passage time by the Lyapunov function method.

On the other hand, the numerical computation of first passage times is usually efficient and accurate. Therefore we choose to check conditions (4)-(7) above numerically when a rigorous proof is out of reach. We do not compute moments directly, because the moments of return times, i.e.,  $\tau^\beta$ , usually do not have finite variances. As a result, a large number of samples is usually necessary to stabilize the estimate of  $\mathbb{E}[\tau^\beta]$ . When the expectation of  $\tau^\beta$  is close to blow-up, numerically verifying whether it is finite becomes even more difficult. Therefore instead, we observe that the finiteness of moments of a random variable is closely related to its tail.

Let  $Z$  be a random variable that takes non-negative integer values. Assume  $\beta > 1$ . The following two lemmas are straightforward.

**Lemma 3.1.** *If  $\mathbb{E}[Z^\beta] < \infty$ , then*

$$\lim_{n \rightarrow \infty} n^\beta \mathbb{P}[Z > n] = 0.$$

*Proof.* First note that,

$$\begin{aligned} \sum_{n=0}^{\infty} n^{\beta-1} \mathbb{P}[Z > n] &= \sum_{n=0}^{\infty} n^{\beta-1} \sum_{m=n+1}^{\infty} \mathbb{P}[Z = m] \\ &= \sum_{m=0}^{\infty} \left( \sum_{n=0}^{m-1} n^{\beta-1} \right) \mathbb{P}[Z = m] \leq C(\beta) \mathbb{E}[Z^{\beta}] < \infty \end{aligned}$$

for some constant  $C(\beta)$ . Then, let  $c_n = \mathbb{P}[Z > n]$  so that  $c_0 \geq c_1 \geq \dots$ . We claim that if

$$\sum_{n=0}^{\infty} c_n n^{\beta-1} < \infty$$

then

$$c_n n^{\beta} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

To see that, observe that

$$\sum_{n=0}^{\infty} (n+1)^{\beta} (c_n - c_{n+1}) = \sum_{n=0}^{\infty} [(n+1)^{\beta} - n^{\beta}] c_n < C_1(\beta) \sum_{n=0}^{\infty} c_n n^{\beta-1}$$

for some constant  $C_1(\beta)$ . So the sum on the left side converges. Since

$$\sum_{n \geq N} (n+1)^{\beta} (c_n - c_{n+1}) = c_N (N+1)^{\beta} + \sum_{n \geq N+1} c_n ((n+1)^{\beta} - n^{\beta}),$$

it follows that both terms on the right tend to 0 as  $N \rightarrow \infty$ .  $\square$

**Lemma 3.2.** *If*

$$\mathbb{P}[Z > n] \leq C n^{-\beta},$$

*then for any  $\epsilon > 0$ , we have*

$$\mathbb{E}[Z^{\beta-\epsilon}] < \infty.$$

*Proof.* For any  $\epsilon > 0$ , we have

$$\begin{aligned} \mathbb{E}[Z^{\beta-\epsilon}] &= \sum_{n=0}^{\infty} n^{\beta-\epsilon} \mathbb{P}[Z = n] \\ &\leq C(\beta) \sum_{n=0}^{\infty} \sum_{m=0}^n m^{\beta-1-\epsilon} \mathbb{P}[Z = n] \\ &\leq C(\beta) \sum_{n=0}^{\infty} (n+1)^{\beta-1-\epsilon} \mathbb{P}[Z > n] \\ &\leq C(\beta) \sum_{n=0}^{\infty} (n+1)^{\beta-1-\epsilon} \min\{1, C n^{-\beta}\} < \infty, \end{aligned}$$

where  $C(\beta)$  is a constant depending on  $\beta$ .  $\square$



Therefore, condition (4) can be checked by computing the tails of  $\tau_{\mathfrak{C}}$ . The polynomial tail of  $\tau_{\mathfrak{C}}$  can be computed by measuring the slope of  $\mathbb{P}[\tau_{\mathfrak{C}} > n]$  vs.  $n$  in the log-log plot. If  $\mathbb{P}[\tau_{\mathfrak{C}} > n] \sim n^{-\beta}$  for some  $\beta > 1$ , then by Lemma 3.2, we have  $\mathbb{E}[\tau_{\mathfrak{C}}^{\beta-\epsilon}] < \infty$  for any small  $\epsilon > 0$ , which implies condition (4) for parameter  $\beta - \epsilon$ . Condition (7) can be verified in the same way by taking the numerical steady-state distribution as the initial condition.

It remains to check conditions (5) and (6). When  $\beta > 1$ , it is easy to see that (6) is implied by (5) immediately. There are two approaches to check condition (5).

- Numerically find the maximum of  $\mathbb{E}_x[\tau_{\mathfrak{C}}^{\beta}]$  for  $x \in \mathfrak{C}$ , denoted by  $x^*$ . Then compute the numerical tail of  $\tau_{\mathfrak{C}}|_{x^*}$ , i.e.,  $\mathbb{P}_{x^*}[\tau_{\mathfrak{C}} > n]$ .
- If  $\mathfrak{C}$  is compact, condition (5) follows from the continuity of  $\mathbb{E}_x[\tau_{\mathfrak{C}}^{\beta}]$  for  $x \in \mathfrak{C}$ . The continuity can be verified either analytically or numerically.

In summary, we are interested in the verification of the following four conditions:

- (A1)  $\Psi_n$  is irreducible with respect to a non-trivial probability measure  $\phi$ .
- (A2)  $\Psi_n$  admits a uniform reference set  $\mathfrak{C}$  such that

$$\mathcal{P}(x, \cdot) \geq \eta \theta(\cdot) \quad \text{for all } x \in \mathfrak{C}$$

and

$$\theta(\mathfrak{C}) > 0.$$

- (N1) Distributions  $\mathbb{P}_{\mu}[\tau_{\mathfrak{C}} \geq n]$  and  $\mathbb{P}_{\pi}[\tau_{\mathfrak{C}} \geq n]$  have polynomial tails  $\sim n^{-\beta}$  for some  $\beta > 1$ , where  $\mu$  is the initial distribution that we are interested in, and  $\pi$  is the numerically obtained invariant measure.
- (N2) The distribution  $\mathbb{P}_{x^*}[\tau_{\mathfrak{C}} \geq n]$  has a polynomial tail  $\sim n^{-\beta}$  for the  $\beta > 1$  above, where  $\mathbb{E}_{x^*}[\tau_{\mathfrak{C}}^{\beta}]$  is the numerical maximum of  $\mathbb{E}_x[\tau_{\mathfrak{C}}^{\beta}]$  for  $x \in \mathfrak{C}$ .

Condition (N2) can be replaced by

(N2)' Function

$$f(x) := \mathbb{E}_x[\tau_{\mathfrak{C}}^{\beta}]$$

is continuous with respect to  $x \in \mathfrak{C}$ .  $\mathfrak{C}$  is a compact set.

*Remark 3.3.* The rigorous result does not guarantee that the speed of convergence equals the speed of contraction of the Markov operator. It is possible that there exists a  $\beta_1$  such that (N2) holds for  $\beta_1$  and there exist initial distributions  $\mu, \nu$  such that

$$\mathbb{E}_{\mu}[\tau_{\mathfrak{C}}^{\beta_1}] < \infty, \quad \text{and} \quad \mathbb{E}_{\nu}[\tau_{\mathfrak{C}}^{\beta_1}] < \infty,$$

but  $\mathbb{E}_{\pi}[\tau_{\mathfrak{C}}^{\beta_1}] = \infty$ . In fact, in this situation the renewal theory only implies  $\mathbb{E}_{\pi}[\tau_{\mathfrak{C}}^{\beta_1-1}] < \infty$ . In other words, the speed of contraction of the Markov operator may be faster than the speed of convergence to the invariant probability measure.

## 4. MAIN CONCLUSION FROM CONDITIONS

This section discusses the main conclusions one can obtain from the sufficient conditions summarized in the previous section. What one can learn from a direct numerical simulation is simple, as it only shows the speed of convergence with respect to one initial distribution. Different from that, our method can support more general results as it incorporates both numerical simulations and analytical proof.

Applying Lemma 3.2 and Theorem 2.5, we can obtain the following direct consequences from (A1), (A2), (N1) and (N2).

- There exists an invariant probability measure  $\pi$ .
- For the initial distribution  $\mu$  we have tested,

$$\lim_{n \rightarrow \infty} n^{\beta-\epsilon} \|\mu \mathcal{P}^n - \pi\|_{TV} = 0.$$

for any  $\epsilon > 0$ .

In this section, we will try to go beyond that.

**4.1. Initial distributions.** First we will show that the sufficient condition in the previous subsection implies the polynomial convergence rate for a wider class of initial conditions. This follows immediately from the following proposition.

**Proposition 4.1.** *Assume  $\Psi_n$  is irreducible with respect to  $\phi$ . If*

$$\sup_{x \in C} \mathbb{E}_x[\tau_C^\beta] < \infty,$$

*Then  $\mathbb{E}_x[\tau_C^\beta] < \infty$  for  $\phi$ -almost every  $x \in X$ .*

*Proof.* Let  $A \in \mathcal{B}$  be the set such that  $\forall x \in A, \mathbb{E}_x[\tau_C^\beta] = \infty$ . Suppose that  $\phi(A) > 0$ . Since  $\Psi_n$  is irreducible,  $\sum_{n=1}^{\infty} P^n(x, A) > 0$  for all  $x \in C$ . Choose  $x_0 \in C$ , by irreducibility there exists  $n > 0$  such that  $P^n(x_0, A) > 0$ .

Define the  $n$ -step taboo transition probability by

$${}_A P^n(n, B) := P_x(\Phi_x \in B, \tau_A \geq n),$$

where  $x \in X, B \in \mathcal{B}$ , to be the probability of a transition to  $B$  in  $n$  steps of the chain, avoiding the set  $A$ . By the following last exit decomposition, we have

$$P^n(x_0, A) = \sum_{k=1}^n \int_C P^k(x_0, dy) {}_C P^{n-k}(y, A) > 0.$$

Therefore, there exists  $x_1 \in C$  and  $m > 0$  such that  ${}_C P^m(x_1, A) > 0$ . This implies

$$\begin{aligned} \mathbb{E}_{x_1}[\tau_C^\beta] &= \sum_{n=0}^{\infty} n^\beta \cdot {}_C P^n(x_1, C) \\ &\geq \sum_{n=m}^{\infty} n^\beta \int_A {}_C P^m(x_1, dw) \cdot {}_C P^{n-m}(w, C) \\ &\geq \int_A {}_C P^m(x_1, dw) \mathbb{E}_w[\tau_C] \\ &= \infty. \end{aligned}$$

But  $\sup_{x \in C} \mathbb{E}_x[\tau_{\mathcal{C}}^\beta] < \infty$  and  $x_1 \in C$  imply that  $\mathbb{E}_{x_1}[\tau_{\mathcal{C}}^\beta] < \infty$ , resulting in a contradiction. Therefore  $\phi(A) = 0$ .  $\square$

By proposition 4.1, if condition **(N2)** holds, then **(N1)** automatically holds for  $\phi$ -almost all  $\mu = \delta_x$ ,  $x \in X$ .

**4.2. Decay of correlation.** Let  $\xi(x)$  and  $\eta(x)$  be two functions in  $L^\infty(X)$ . The decay of correlation  $C_\mu^{\xi, \eta}(n)$  is denoted by

$$C_\mu^{\xi, \eta}(n) := \left| \int (\mathcal{P}^n \eta)(x) \xi(x) \mu(dx) - \int (\mathcal{P}^n \eta)(x) \mu(dx) \int \xi(x) \mu(dx) \right|.$$

We have the following proposition regarding the decay of correlation.

**Proposition 4.2.** *If **(A1)**, **(A2)**, **(N1)** and **(N2)** are satisfied, then*

$$C_\mu^{\xi, \eta}(n) \leq O(1) \cdot n^{\epsilon - \beta}$$

for any  $\epsilon > 0$ .

*Proof.* We have

$$\begin{aligned} & \left| \int (P^n \eta)(\mathbf{E}) \xi(\mathbf{E}) \mu(d\mathbf{E}) - \int (P^n \eta)(\mathbf{E}) \mu(d\mathbf{E}) \int \xi(\mathbf{E}) \mu(d\mathbf{E}) \right| \\ &= \left| \int \xi(\mathbf{E}) \left( (P^n \eta)(\mathbf{E}) - \int (P^n \eta)(\mathbf{Z}) \mu(d\mathbf{Z}) \right) \mu(d\mathbf{E}) \right| \\ &\leq \|\xi\|_{L^\infty} \|\eta\|_{L^\infty} \int \|\delta_{\mathbf{E}} P^n - \mu P^n\|_{TV} \mu(d\mathbf{E}). \end{aligned}$$

Since **(A1)**, **(A2)**, **(N1)** and **(N2)** hold, it follows from Corollary 2.6 that

$$\|\delta_{\mathbf{E}} P^n - \mu P^n\|_{TV} \leq C(\mathbf{E}_{\mathbf{E}}[\tau_{\mathcal{C}}^{\beta - \epsilon}] + \mathbf{E}_{\mu}[\tau_{\mathcal{C}}^{\beta - \epsilon}]) n^{\epsilon - \beta}.$$

We have  $\mathbf{E}_{\mu}[\tau_{\mathcal{C}}^{\beta - \epsilon}] < \infty$  and

$$\int \mathbf{E}_{\mathbf{E}}[\tau_{\mathcal{C}}^{\beta - \epsilon}] \mu(d\mathbf{E}) = \mathbf{E}_{\mu}[\tau_{\mathcal{C}}^{\beta - \epsilon}] < \infty.$$

Therefore,

$$\|\delta_{\mathbf{E}} P^n - \mu P^n\|_{TV} \leq O(1) \cdot n^{\epsilon - \beta}.$$

This completes the proof.  $\square$

**4.3. Results for time-continuous process.** We have only showed sufficient conditions of polynomial convergence rate for time-discrete Markov processes. If  $\Psi_t$  is a time-continuous Markov process, our method is also applicable after some additional effort. Instead of investigating the infinitesimal generator of  $\Psi_t$ , we will work on the time- $h$  sample chain of  $\Psi_t$ , which is denoted by  $\Psi_n^h := \Psi_{nh}$ . We remark that certain drift conditions about the infinitesimal generator of  $\Psi_t$  also implies the polynomial ergodicity [11]. However, these drift conditions require a Lyapunov function, which is not explicitly known in general. Our argument about the tail of the first return time is more suitable for a numerical investigation.

To pass the existence of an invariant measure from  $\Psi_n^h$  to  $\Psi_t$ , the “continuity at zero” is necessary. We will show that in the following proposition.

**Proposition 4.3.** *Assume  $\Psi_n^h$  is aperiodic and admits an invariant probability measure  $\pi_h$ . If in addition,*

$$\|\pi_h P^\delta - \pi_h\|_{TV} \rightarrow 0 \quad \text{as } \delta \rightarrow 0,$$

*then  $\pi = \pi_h$  is invariant for any  $\Psi_t$ .*

*Proof.* Notice that  $\pi_h$  is invariant for any  $\Psi_n^{hj/k}$ , where  $j, k \in \mathbb{Z}^+$  (Theorem 10.4.5 of [21]). Then without loss of generality, assume  $t/h \notin \mathbb{Q}$ . By the density of orbits in irrational rotations, there exists sequences  $a_n, b_n \in \mathbb{Z}^+$  such that

$$d_n := t - \frac{a_n}{b_n}h \rightarrow 0$$

from right. Then

$$\pi_h P^t = \pi_h P^{\frac{a_n}{b_n}h} P^{d_n}.$$

Therefore,

$$\|\pi_h P^t - \pi_h\|_{TV} \leq \lim_{n \rightarrow \infty} \|\pi_h P^{d_n} - \pi_h\|_{TV} = 0$$

by the assumption about the “continuity at zero”. Hence  $\pi_h$  is invariant with respect to  $P^t$ .  $\square$

**Theorem 4.4.** *Assume (A1), (A2), (N1) and (N2) hold for  $\Psi_n^h$ . If in addition*

$$\|\pi_h P^\delta - \pi_h\|_{TV} \rightarrow 0 \quad \text{as } \delta \rightarrow 0,$$

*where  $\pi_h$  is the invariant measure for  $\Psi_n^h$ , then for any small  $\epsilon > 0$ ,*

$$\lim_{t \rightarrow \infty} t^{\beta-\epsilon} \|\mu P^t - \pi_h\|_{TV} = 0.$$

*Proof.* It follows from Proposition 4.3 that  $\pi_h$  is invariant with respect to  $P^t$  for any  $t > 0$ . In addition,

$$\|\mu P^t - \pi_h\|_{TV} = \|\mu P^{nh} P^r - \pi_h\|_{TV} = \|(\mu P^{nh} - \pi_h) P^r\|_{TV} \leq \|\mu P^{nh} - \pi_h\|_{TV},$$

where  $n$  is the greatest integer that is smaller than  $t/h$ . This implies

$$\lim_{t \rightarrow \infty} t^{\beta-\epsilon} \|\mu P^t - \pi_h\|_{TV} = 0.$$

$\square$

In many situations, especially when the time step has to be small, it may be easier to simulate  $\tau_{\mathfrak{C}}$  for  $\Psi_t$  instead of  $\Psi_n^h$ . Let  $\hat{\tau}_{\mathfrak{C}}$  be the first-passage time to  $\mathfrak{C}$  for  $\Psi_n^h$ . Let  $h$  be a fixed parameter. The first return time and first passage time for  $\Psi_t$  is defined as

$$\tau_{\mathfrak{C}} = \tau_{\mathfrak{C}}(h) = \inf_{t \geq h} \{\Psi_t \in \mathfrak{C}\} \text{ and } \sigma_{\mathfrak{C}} = \inf_{t \geq 0} \{\Psi_t \in \mathfrak{C}\}.$$

We drop the notation  $(h)$  when it does not lead to confusion. Here we have to treat return time and passage time differently because otherwise  $\sigma_{\mathfrak{C}}$  is zero for all initial conditions within  $\mathfrak{C}$ . We say (N1) and (N2) hold for  $\Psi_t$  with parameters  $h$  and  $\beta$  if

(N1) Distributions  $\mathbb{P}_\mu[\tau_{\mathfrak{C}}(h) \geq n]$  and  $\mathbb{P}_\pi[\tau_{\mathfrak{C}}(h) \geq n]$  have polynomial tails  $\sim n^{-\beta}$  for some  $\beta > 1$ .

(N2) The distribution  $\mathbb{P}_{x^*}[\tau_{\mathfrak{C}}(h) \geq n]$  has a polynomial tail  $\sim n^{-\beta}$  for the  $\beta > 1$  above, where  $\mathbb{E}_{x^*}[\tau_{\mathfrak{C}}(h)^\beta]$  is the numerical maximum of  $\mathbb{E}_x[\tau_{\mathfrak{C}}(h)^\beta]$  for  $x \in \mathfrak{C}$ .

The following theorem gives the relation between tail of  $\hat{\tau}_{\mathfrak{C}}$  and tail of  $\tau_{\mathfrak{C}}$ .

**Proposition 4.5.** *Assume (N1) and (N2) hold for  $\Psi_t$  with parameter  $h$  and  $\beta$ . If further*

$$\inf_{x \in \mathfrak{C}} \mathbb{P}_x[\Psi_h = \Psi_0] > \gamma > 0$$

for the step size  $h$ , then for any  $\epsilon > 0$ , (N1) and (N2) hold for  $\Psi_n^h$  with parameter  $\beta - \epsilon$ .

*Proof.* Define the following stopping times  $\tau_n$  and random times  $\xi_n$ .  $\xi_0 = 1$ .  $\tau_n = \inf_{t \geq \xi_{n-1}h} \{\Psi_t \in \mathfrak{C}\}$ .  $\xi_n = \lceil \tau_n/h \rceil$ . It is possible that  $\Psi_t$  leaves  $\mathfrak{C}$  after  $\tau_n$  but before  $\xi_n$  and thus makes a “false return”. To verify (N1) and (N2) for  $\Psi_n^h$ , we need to estimate the number of “false returns”.

Let  $N$  be the number of “false returns” of  $\Psi_t$ :

$$N = \inf_{n > 0} \{\Psi_{\xi_n h} \in \mathfrak{C}\}.$$

It is easy to see that  $\hat{\tau}_{\mathfrak{C}} = \xi_N$ . Therefore, for any small  $\sigma > 0$ , we have

$$\{\xi_N > n^{1+\delta}\} \subset \{N > n^\delta\} \cup \bigcup_{k=0}^{\lfloor n^\delta \rfloor} \{\xi_{k+1} - \xi_k > n, N > k\}.$$

Without loss of generality we assume  $n > 1$  in the tail estimates throughout the proof. For any  $\Psi_0$ , by the Markov property, we have

$$\mathbb{P}_{\Psi_0}[N = n | N > n-1] \geq \mathbb{P}_{\Psi_{\tau_n}}[\Psi_{\xi_n h} = \Psi_{\tau_n}].$$

Since  $\Psi_{\tau_n} \in \mathfrak{C}$ , this probability is at least  $\delta$ . Therefore, we have

$$\mathbb{P}_{\Psi_0}[N > n] \leq (1 - \gamma)^n$$

for any  $\Psi_0$  and  $n > 1$ .

Let  $\epsilon > 0$  be an arbitrary small number. For  $k \geq 1$ , we have

$$\begin{aligned} \mathbb{P}_{\Psi_0}[\xi_{k+1} - \xi_k > n, N > k] &\leq \mathbb{P}_{\Psi_0}[\tau_{k+1} - \xi_k h > (n-1)h, N > k] \\ &\leq \mathbb{P}_{\Psi_0}[\tau_{k+1} - \xi_k h > (n-1)h, \Psi_{\xi_k h} \notin \mathfrak{C}] \\ &= \mathbb{P}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2} > (n-1)^{\beta-\epsilon/2} h^{\beta-\epsilon/2}, \Psi_{\xi_k h} \notin \mathfrak{C}] \\ &\leq \frac{\mathbb{E}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2} \mathbf{1}_{\Psi_{\xi_k h} \in \mathfrak{C}}]}{(n-1)^{\beta-\epsilon/2} h^{\beta-\epsilon/2}} = \frac{\mathbb{E}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2}]}{(n-1)^{\beta-\epsilon/2} h^{\beta-\epsilon/2}}, \end{aligned}$$

where  $\mathbb{P}_{\Psi_{\xi_k h}}$  and  $\mathbb{E}_{\Psi_{\xi_k h}}$  mean letting the initial distribution to be  $\Psi_{\xi_k h}$ . The second to last inequality follows from Markov inequality. Then it is sufficient to show that  $\mathbb{E}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2}] < \infty$ .

Define  $\tau_{\mathfrak{C},r} = \inf_{t > r} \{\Psi_t \in \mathfrak{C}\} - r$  be the first return time when starting from  $\Psi_r$ . Then

$$\mathbb{E}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2}] = \int_{\mathfrak{C}} \int_0^h \mathbb{E}_x[\tau_{\mathfrak{C},r}^{\beta-\epsilon/2}] \mathbb{P}_{\Psi_0}[\xi_k h - \tau_k = r, \Psi_{\tau_k} = x] dr dx.$$

Notice that  $r \leq h$ . For each sample path starting from  $\tau_k$ , we have  $\tau_{\mathfrak{C},r} \leq \tau_{\mathfrak{C}}(h)$ . This implies

$$\mathbb{E}_{\Psi_{\xi_k h}}[\sigma_{\mathfrak{C}}^{\beta-\epsilon/2}] \leq \int_{\mathfrak{C}} \int_0^h \mathbb{E}_x[\tau_{\mathfrak{C}}^{\beta-\epsilon/2}] \mathbb{P}_{\Psi_0}[\xi_k h - \tau_k = r, \Psi_{\tau_k} = x] dr dx \leq \sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}(h)^{\beta-\epsilon/2}].$$

Since **(N2)** holds for  $\Psi_t$ , by Lemma 3.2, we have

$$\sup_{x \in \mathfrak{C}} \mathbb{E}_x[\tau_{\mathfrak{C}}(h)^{\beta-\epsilon/2}] \leq C_0$$

for some constant  $C_0$ . This implies

$$\mathbb{P}_{\Psi_0}[\xi_{k+1} - \xi_k > n, N > k] \leq C n^{-(\beta-\epsilon/2)}$$

for some constant  $C$ .

If  $k = 0$ , we need to estimate

$$\mathbb{P}_{\Psi_0}[\tau_{\mathfrak{C}}(h) > (n-1)h]$$

for  $\Psi_0 = x^*$ ,  $\Psi_0 \sim \mu$ , and  $\Psi_0 \sim \pi$ , where  $\mathbb{E}_{x^*}[\hat{\tau}_{\mathfrak{C}}^{\beta}]$  is the maximum of  $\mathbb{E}_x[\hat{\tau}_{\mathfrak{C}}^{\beta}]$  for  $x \in \mathfrak{C}$ . Since **(N1)** and **(N2)** hold for  $\Psi_t$ , we have

$$\mathbb{P}_{\Psi_0}[\tau_{\mathfrak{C}}(h) > (n-1)h] \leq C' n^{-\beta} \leq C' n^{-(\beta-\epsilon/2)}$$

for all three cases, where  $C'$  depends on the initial condition.

Therefore, for all three cases  $\Psi_0 = x^*$ ,  $\Psi_0 \sim \mu$ , and  $\Psi_0 \sim \pi$ , we have

$$\begin{aligned} \mathbb{P}_{\Psi_0}[\xi_N > n^{1+\delta}] &\leq \mathbb{P}[N > n^{\delta}] + \sum_{k=0}^{\lfloor n^{\delta} \rfloor} \mathbb{P}_{\Psi_0}[\xi_{k+1} - \xi_k > n, N > k] \\ &\leq (1-\gamma)^{n^{\delta}} + n^{\delta} \cdot \max\{C, C'\} n^{-(\beta-\epsilon/2)}. \end{aligned}$$

Note that  $(1-\gamma)^{n^{\delta}}$  converges to 0 faster than  $n^{-\beta-\epsilon/2}$ . For any small  $\epsilon > 0$ , by making  $\delta$  sufficiently small, we have

$$\mathbb{P}_{\Psi_0}[\hat{\tau}_{\mathfrak{C}} > n] \leq C_1 n^{-(\beta-\epsilon)}$$

for some constant  $C_1$  that depends on the initial distribution. **(N1)** for  $\Psi_n^h$  is verified by cases  $\Psi_0 \sim \mu$  and  $\Psi_0 \sim \pi$ . **(N2)** for  $\Psi_n^h$  is verified by case  $\Psi_0 = x^*$ . This completes the proof.  $\square$

Finally, it is trivial to show that if **(A1)**, **(A2)**, **(N1)** and **(N2)** hold for  $\Psi_n^h$ , then Proposition 4.2 remains true for  $\Psi_t$ .

**4.4. Summary of conclusions.** In summary, assume **(A1)**, **(A2)**, **(N1)** and **(N2)** hold for  $\Psi_n$ , then

- (a)  $\Psi_n$  admits an invariant probability measure  $\pi$ .
- (b) Polynomial convergence rate to  $\pi$ :

$$\lim_{n \rightarrow \infty} n^{\beta-\epsilon} \|\mu \mathcal{P}^n - \pi\|_{TV} = 0$$

for any  $\epsilon > 0$ .

(c) Polynomial decay rate of correlation:

$$\lim_{n \rightarrow \infty} n^{\beta-\epsilon} C_{\pi}^{\xi, \eta} = 0$$

for any  $\epsilon > 0$ .

(d) Polynomial convergence rate to  $\pi$ . for any  $\epsilon > 0$ , we have

$$\lim_{n \rightarrow \infty} n^{\beta-\epsilon} \|\delta_x \mathcal{P}^n - \pi\|_{TV} = 0$$

for  $\phi$ -almost every  $x \in X$ .

If  $\Psi_t$  is a time-continuous Markov process with transition kernel  $\mathcal{P}^t$ . Assuming that the “continuity at zero” condition in Proposition 4.3 is satisfied, if **(A1)**, **(A2)**, **(N1)** and **(N2)** hold for  $\Psi_n^h$ , then conclusions (a)-(d) also hold for  $\Psi_t$ . If **(N1)** and **(N2)** hold for  $\Psi_t$ , then conclusions (a)-(d) still hold for  $\Psi_t$  as one can put  $\epsilon/2$  into Proposition 4.5 and Theorem 2.5.

## 5. EXAMPLE: STOCHASTIC ENERGY EXCHANGE MODEL

**5.1. Derivation from deterministic dynamics.** Consider a long tube of gas that is connected to two thermalized boundaries. Assume that all gas molecules are rigid moving particles and that the only interaction between particles is rigid body collision. Then we have a very complicated deterministic dynamical system like as described in Figure 1. Apparently this is a very difficult multi-body problem. To reduce its significant difficulty, one approach is to “localize” all particles such that particles are trapped in one-dimensional cells, like the one described in Figure 2. This is called the locally confined particle system, in which energy transport still exists because neighboring particles can collide through the gates between cells.

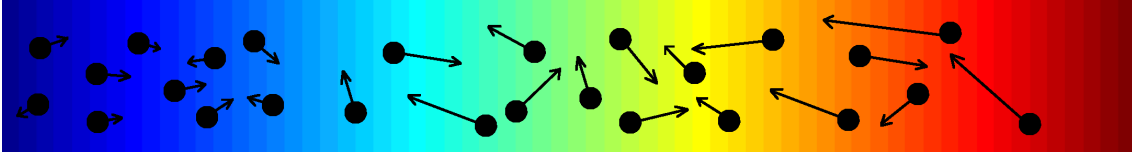


FIGURE 1. Moving particles in a long tube.

As probably the simplest deterministic microscopic heat conduction model, the locally confined particle system is studied in various literatures. For example, under certain assumptions, the ergodicity of this model has been proved in [3]. Further rigorous investigation of the locally confined particle system is known to be difficult. Each cell in the locally confined particle system is a chaotic billiard table. And chaotic billiards are known to have many stochastic properties. Therefore, a natural strategy is to use a Markov process to approximate the change of particle energy in the locally confined particle system.

In [15], numerical simulations of first particle-particle collision in the locally confined particle system shows that for a pair of adjacent particles with energy  $(E_1, E_2)$ , the first particle-particle collision time has an exponential tail with slope  $R(E_1, E_2)$ . In addition, when  $\min\{E_1, E_2\} \ll 1$ , we have  $R(E_1, E_2) \sim \sqrt{\min\{E_1, E_2\}}$ . The

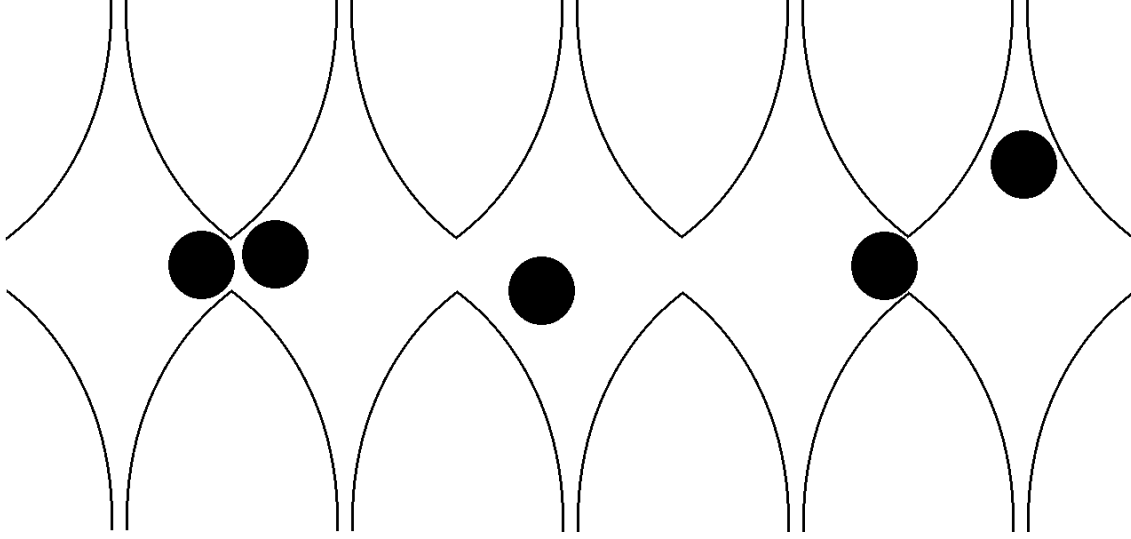


FIGURE 2. Particles are confined in a chain of cells.

heuristic justification of this rate is that when a particle has low energy, it has to move to the gate between cells by itself in order to have the next energy exchange. Hence a very slow particle dominates the next particle-particle collision time. If in addition we assume energy exchange in a particle-particle collision is done in a “random halves” way, we have a stochastic energy exchange model described in the next subsection.

**5.2. Model Description.** Consider a chain of  $N$  linearly ordered lattice sites  $\{1, 2, \dots, N\}$ , each storing a fixed amount of energy  $E_i$ ,  $i = 1, 2, \dots, N$ . The chain is connected to two heat baths at the ends with temperatures  $T_L$  and  $T_R$ , respectively.

An exponential clock is associated with each pair of adjacent sites, with rate  $R(E_i, E_{i+1}) = \sqrt{\min\{E_i, E_{i+1}\}}$ , called the stochastic energy exchange rate. Two clocks are associated with the ends of the chain and the heat baths. The clock between the left (resp. right) heat bath and the first (resp. last) site has a rate  $\sqrt{\min\{T_L, E_1\}}$  (resp.  $\sqrt{\min\{T_R, E_N\}}$ ).

When the  $i$ -th clock rings, sites  $i$  and  $i + 1$  exchange energy as

$$(E'_i, E'_{i+1}) = (p(E_i + E_{i+1}), (1 - p)(E_i + E_{i+1})),$$

where  $p$  satisfies the uniform distribution on  $(0, 1)$  and is independent of everything else. When a clock involving heat bath rings, the corresponding site exchanges energy with an exponential random variable with mean  $T_L$  (or  $T_R$ ) in the same “random halves” fashion. This is to say, we have

$$E'_1 = p(E_1 + e_L), \quad e_L \sim \text{Exp}(T_L)$$

and

$$E'_N = p(E_N + e_R), \quad e_R \sim \text{Exp}(T_R),$$



where  $p$  still satisfies the uniform distribution on  $(0, 1)$  and is independent of everything else. For the sake of simplicity, clocks at left and right boundaries are denoted by clock 0 and clock  $N + 1$ , respectively.

It is easy to see that the stochastic energy exchange model generates a Markov jump process  $\mathbf{E}_t = (E_1(t), \dots, E_N(t))$  on  $\mathbb{R}_+^N$ . We denote the transition kernel of  $\mathbf{E}_t$  by  $P^t(\mathbf{E}, \cdot)$  for  $\mathbf{E} \in \mathbb{R}_+^N$ .

Let  $h > 0$  be a fixed number that represents the step size. The time- $h$  sampling chain of  $\mathbf{E}_t$  is denoted by  $\mathbf{E}_n^h$ , or simply  $\mathbf{E}_n$  when it does not lead to a confusion. The transition kernel of  $\mathbf{E}_n$  is denoted by  $P(\mathbf{E}, \cdot)$ .

**5.3. Verifying (A1) and (A2).** We will first work on the time- $h$  chain  $\mathbf{E}_n$ . The verification of analytical conditions of  $\mathbf{E}_n$  is based on the following Theorem.

**Theorem 5.1.** *For any set  $K \subset \mathbb{R}_+^N$  of the form  $K = \{(E_1, \dots, E_N) \mid 0 < c_i \leq E_i \leq C_i, i = 1 \sim N\}$  and any  $h > 0$ , there exists a constant  $\eta > 0$  such that*

$$P(\mathbf{E}, \cdot) > \eta U_K(\cdot),$$

for any  $\mathbf{E} \in K$ , where  $U_K$  is the uniform probability distribution over  $K$ .

*Proof.* For  $\mathbf{E} = \{E_1, \dots, E_N\} \in K$  and  $d\mathbf{E} = \{(dE_1, \dots, dE_N), dE_i > 0, i = 1 \sim N\}$ . Assume  $dE_i$ s are sufficiently small. Let

$$B(\mathbf{E}, d\mathbf{E}) = \{(x_1, \dots, x_N) \in \mathbb{R}^N \mid E_i \leq x_i \leq E_i + dE_i\}$$

denote the hypercube in  $\mathbb{R}^N$ . It then suffices to prove that for any  $\mathbf{E}_0 = \{\bar{E}_1, \dots, \bar{E}_N\} \in K$ , we have

$$(5.1) \quad P(\mathbf{E}_0, B(\mathbf{E}, d\mathbf{E})) > \sigma dE_1 dE_2 \cdots dE_N,$$

where  $\sigma$  is a strictly positive constant that is independent of  $\mathbf{E}$  and  $\mathbf{E}_0$ .

We will then construct a sequence of events to go from the state  $\mathbf{E}_0$  to  $B(\mathbf{E}, d\mathbf{E})$  with desired positive probability. Let  $\delta = \frac{h}{2N+1}$  and let  $\epsilon > 0$  be sufficiently small such that  $\epsilon < \min\{c_i, i = 1 \sim N\}$ . Let  $H = \sum_{i=1}^N E_i + dE_i$ . We consider the events  $S_1 \cdots, S_N$  and  $F_1, \dots, F_{N+1}$ , where  $S_i$  and  $F_j$  specifies what happens on the time interval  $(i\delta, (i+1)\delta]$  and  $(N\delta + (j-1)\delta, N\delta + j\delta]$ , respectively.

- $S_i = \{E_i(i\delta) \in [\epsilon/2, \epsilon]\}$  and  $\{\text{the } i\text{-th clock rings exactly once, all other clocks are silent on } ((i-1)\delta, i\delta]\}$ .
- $F_1 = \{\text{Energy emitted by right heat bath} \in (H, 2H)\}$  and the  $N$ -th clock rings exactly once, all other clocks are silent on  $(N\delta, (N+1)\delta]$ .
- $F_j = \{E_j(N\delta + j\delta) \in [E_{N+2-j}, E_{N+2-j} + dE_j]\}$  and  $\{\text{the } (N+1-j)\text{-th clock rings exactly once, all other clocks are silent on } (N\delta + (j-1)\delta, N\delta + j\delta]\}$  for  $j = 2, \dots, N+1$ .

The idea is that the energy at each site is first emptied to the right heat bath. Then a sufficiently large amount of energy is injected into the chain from the right heat bath so that it is always possible for site  $j$  to acquire an amount of energy between  $E_j$  and  $E_j + dE_j$  by passing the rest to site  $j-1$ , where sites 0 and  $N+1$  denote the left and right heat baths respectively.

It is easy to show that the probability of occurrence of the sequence of events described above is always strictly positive. Here is a brief list of considerations. We will leave detailed calculations to the reader.

- (a) At each clock tick, we can give a lower bound on the rate of the  $i^{\text{th}}$  clock by  $\sqrt{\min\{E_i, E_{i+1}\}} > 0$  since  $0 < c_i \leq E_i$  for all  $i = 1 \sim N$ .
- (b) Let  $p \in (0, 1)$  be the fraction in the mixing that puts  $E_i \in [\epsilon/2, \epsilon]$ . From the rule of energy redistribution, we need  $\epsilon/2 \leq p(E_i + E_{i+1}) \leq \epsilon$ . Rearrange the terms to get  $\epsilon/[2(E_i + E_{i+1})] \leq p \leq \epsilon/(E_i + E_{i+1})$ . This is possible since  $c_i + c_{i+1} \leq E_i + E_{i+1} \leq C_i + C_{i+1}$ . Hence probabilities of  $S_i$  are strictly positive.
- (c) There is also a uniform upper bound on  $H$  given by  $2 \sum_{i=1}^N C_i$ .
- (d) Let  $p \in (0, 1)$  be the fraction in the mixing that puts  $E_j \in (E_j, E_j + dE_j)$ . From the rule of energy redistribution, we have  $E_j/(E_{j-1} + E_j) \leq p \leq (E_j + dE_j)/(E_{j-1} + E_j)$ . Because of  $c_{j-1} + c_j \leq E_{j-1} + E_j \leq C_{j-1} + C_j$  and  $dE_j > 0$ ,  $\mathbb{P}[E_j(N\delta + j\delta) \in [E_{N+2-j}, E_{N+2-j} + dE_j]] > \alpha dE_j$  for some strictly positive constant  $\alpha$ . Hence probabilities of  $F_j$  are greater than  $\text{const} \cdot dE_j$ .

In addition, all these probabilities are uniformly bounded from below for all  $\mathbf{E}$  and  $\bar{\mathbf{E}}$  in  $K$ . Hence we have

$$\mathbb{P}[S_1 \cdots S_N F_1 \cdots F_{N+1}] \geq \sigma dE_1 \cdots dE_N$$

for some constant  $\sigma > 0$ . □

As a corollary, we can prove that  $\mathbf{E}_n$  is both aperiodic and irreducible with respect to the Lebesgue measure.

**Corollary 5.2.**  $\mathbf{E}_n$  is a strongly aperiodic Markov chain.

*Proof.* By theorem 5.1,  $K$  is a uniform reference set. In addition  $U_K(K) > 0$ . The strong aperiodicity follows from its definition. □

Therefore  $\mathbf{E}_n$  is aperiodic.

**Corollary 5.3.**  $\mathbf{E}_n$  is  $\lambda$ -irreducible, where  $\lambda$  is the Lebesgue measure on  $\mathbb{R}_+^N$ .

*Proof.* Let  $A \subset \mathbb{R}_+^N$  be a set with strictly positive Lebesgue measure. Then there exists a set  $K$  that has the form  $\{(E_1, \dots, E_N) \mid 0 < c_i \leq E_i \leq C_i, i = 1 \sim N\}$  and  $U_K(K \cap A) > 0$ .

For any  $\mathbf{E}_0 \in \mathbb{R}_+^N$  and the time step  $h > 0$ , we can choose a  $K \subset \mathbb{R}_+^N$  of the form  $K = \{(E_1, \dots, E_N) \mid 0 < c_i \leq E_i \leq C_i, i = 1 \sim N\}$  for some  $c_i > 0$  and  $C_i < \infty$ , such that  $\mathbf{E}_0 \in K$ . Same construction as in Theorem 5.1 implies that  $P^h(\mathbf{E}_0, \cdot) > \eta U_K(\cdot)$  for some  $\eta > 0$ . Therefore,  $P^h(\mathbf{E}_0, A) > \eta U_K(A) > 0$ . □

Hence assumption **(A1)** and **(A2)** are satisfied.

**5.4. Absolute continuity of invariant measure.** This subsection aims to prove the absolute continuity of  $\pi$  with respect to the Lebesgue measure, which is denoted by  $\lambda$ .

**Proposition 5.4.** *If  $\pi$  is an invariant measure of  $\mathbf{E}_t$ , then  $\pi$  is absolutely continuous with respect to  $\lambda$  with a strictly positive density.*

For  $\mathbf{E} \in \mathbb{R}_+^N$  and  $t > 0$ , we have decomposition

$$P^t(\mathbf{E}, \cdot) = \nu_\perp + \nu_{abs},$$

where  $\nu_{abs}$  and  $\nu_\perp$  are absolutely continuous and singular component with respect to  $\lambda$ , respectively. We need to show that an absolutely continuous component cannot revert back to singularity as time evolves.

**Lemma 5.5.** *For any probability measure  $\mu \ll \lambda$ ,  $\mu P^t \ll \lambda$  for any  $t > 0$ .*

*Proof.* This proof is essentially identical to that of Lemma 6.3 of [17]. All what we need to prove is that, for any absolutely continuous initial distribution, the push-forward measure corresponding to one clock ringing is still absolutely continuous. We refer readers to [17] for the detailed calculations.  $\square$

**Proof of Proposition 5.4.** Let  $\pi = \pi_{abs} + \pi_{perp}$  be an invariant measure. Assume  $\pi_{perp} \neq 0$ . For  $t > 0$ ,  $\pi_{abs} P^t \ll \lambda$  by Lemma 5.5. By Theorem 5.1, for any  $\mathbf{E} \in \mathfrak{C}$ ,  $P^{t/2}(\mathbf{E}, \cdot)$  has a strictly positive density on  $\mathfrak{C}$ . Since  $\mathfrak{C}$  is accessible within finite many energy exchanges,  $P^{t/2}(\mathbf{E}, \mathfrak{C}) > 0$  for all  $\mathbf{E} \in \mathbb{R}_+^N$ . Hence  $P^t(\mathbf{E}, \cdot)$  has a strictly positive density on  $\mathfrak{C}$  for all  $\mathbf{E} \in \mathbb{R}_+^N$ . Therefore,  $\pi_\perp P^t$  has an absolutely continuous component. The absolutely continuous component of  $\pi P^t$  is strictly large than that of  $\pi$ . This contradicts with the invariance of  $\pi$ .  $\square$

**5.5. Verifying (N1) and (N2).** Now we are ready to present our numerical results. We let  $N = 3$  or  $4$  depending on the computational cost of the simulation. The uniform reference set  $\mathfrak{C}$  is chosen as

$$\mathfrak{C} = \{(E_1, \dots, E_N) \mid 0.1 \leq E_i \leq 100, i = 1 \sim N\}.$$

Throughout our numerical justification, we let  $h = 0.1$ . (Recall that for a time-continuous Markov process  $\Psi_t$ , the definition of  $\tau_{\mathfrak{C}} = \tau_{\mathfrak{C}}(h)$  depends on  $h$ .) We will verify (N1) for the numerical invariant measure, which is generated by running the process for a sufficiently long time from a suitable initial distribution. In our simulation, the initial distribution for the simulation of the numerical invariant measure is  $\mu_0 \sim (X_1, \dots, X_N)$ , where  $X_i$  is an exponential distribution with mean  $(T_L + T_R)/2$ . We find that when  $T = 200$ , expectations of many observables we have tested are stabilized (Figure 3). Therefore, the numerical invariant measure is chosen as  $\mu_0 P^{200}$ . Then from the result of our simulation that chooses the numerical invariant measure as the specific starting state, we conclude that  $\mathbb{P}_\pi[\tau_{\mathfrak{C}} > t] \sim t^{-2}$ . (Figure 4).

It remains to check (N2). Instead of numerically showing the continuity of  $f(\mathbf{E}) := \mathbf{E}_{\mathbf{E}}[\tau_{\mathfrak{C}}^{2-\epsilon}]$ , we look for the maximum of  $f(\mathbf{E})$ . Figure 5 shows that when  $\epsilon = 0.3$ ,  $f(\mathbf{E})$  decreases monotonically with decreasing site energy. Therefore, we believe that the maximal of  $f(\mathbf{E})$  in  $\mathfrak{C}$  is  $\mathbf{E}_* = (0.1, 0.1, 0.1)$ . In fact, intuitively one should expect  $f(\mathbf{E})$  to decrease with site energy as starting from low site energy means having higher probability to have even lower site energy after an energy

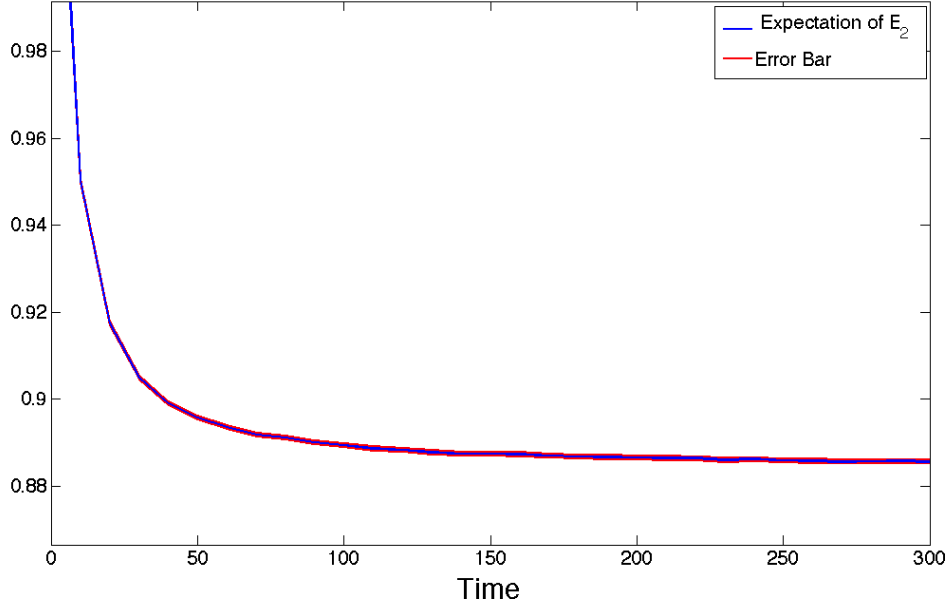


FIGURE 3. Expectation of energy at site 2 verses time. Sample size  $= 1 \times 10^8$ .

exchange. We note that we didn't use numerical optimization tools to find  $\mathbf{E}^*$  due to the low accuracy of  $f(\mathbf{E})$ . When  $\epsilon$  is small,  $\tau_{\mathcal{E}}^{2-\epsilon}$  does not have a finite variance. The Monte Carlo simulation of  $\tau_{\mathcal{E}}^{2-\epsilon}$  inevitably has significant numerical error.

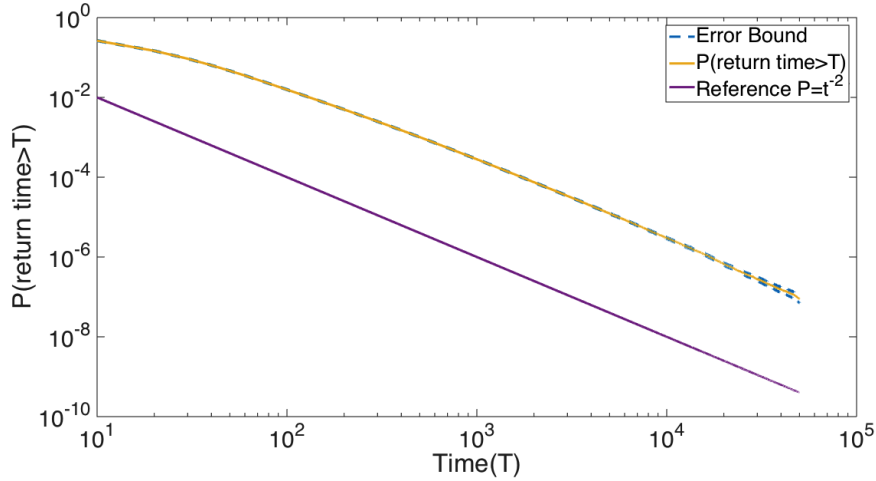


FIGURE 4.  $\mathbb{P}_{\pi}[\tau_{\mathcal{E}} > t]$ ,  $\pi$  is the numerical steady state obtained from long time averaging. Sample size  $= 1 \times 10^8$ . The purple line is a reference line with slope  $-2$ .

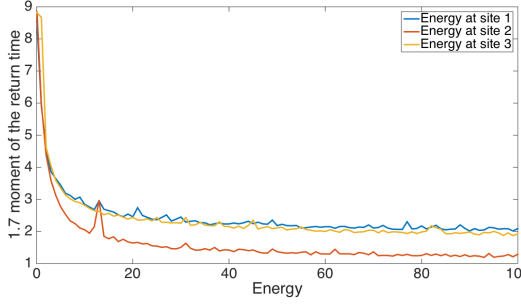


FIGURE 5. Change of  $\mathbb{E}_{\mathbf{E}}[\tau_{\mathcal{C}}^{1.7}]$  for varying  $\mathbf{E}$  when one coordinate of  $\mathbf{E}$  changes. Sample size  $= 1 \times 10^8$  for each initial condition.

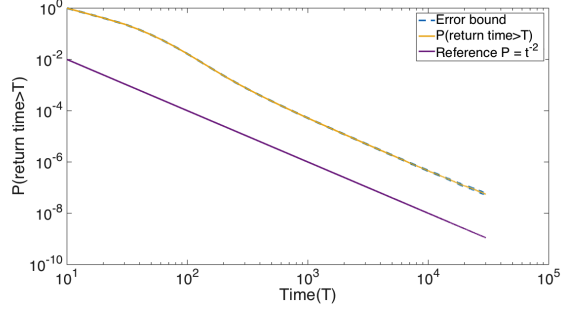


FIGURE 6.  $\mathbb{P}_{\mathbf{E}_*}[\tau_{\mathcal{C}} > t]$  for  $\mathbf{E}_* = (0.1, 0.1, 0.1)$ . Sample size  $= 1 \times 10^{10}$ . The purple line is a reference line with slope  $-2$ .

**5.6. Main conclusions.** The previous subsection verifies both **(N1)** and **(N2)** for  $\mathbf{E}_t$  with parameter 2. The slope of  $\mathbb{P}_{\pi}[\tau_{\mathcal{C}} > t]$  in the log-log plot is 2. Note that  $\mathbb{P}[\Psi_t = \Psi_0 | \Psi_0 \in K] = \mathbb{P}_{\Psi_0}[\text{no clock rings up to } t]$  is uniformly positive for each given  $t > 0$ . By Proposition 4.5, **(N1)** and **(N2)** hold for  $\mathbf{E}_n$  with parameter  $2 - \epsilon/2$  for arbitrarily small  $\epsilon > 0$ . Therefore, conclusions (a)-(d) in Section 4.4 holds for  $\mathbf{E}_n$ .

It remains to pass the results for  $\mathbf{E}_n$  to  $\mathbf{E}_t$ . By Proposition 4.3, it is sufficient to prove “continuity at zero” for  $\mathbf{E}_t$ .

**Lemma 5.6.** *For any probability measure  $\mu$  on  $\mathbb{R}_+^N$ ,*

$$\lim_{\delta \rightarrow 0} \|\mu P^\delta - \mu\|_{TV} = 0$$

*Proof.* It is sufficient to prove that for any  $\epsilon > 0$ , there exists a  $\delta > 0$  such that

$$\|\mu P^\delta - \mu\|_{TV} \leq \epsilon.$$

Since  $\mu$  is finite, there exists a bounded set  $A \subset \mathbb{R}_+^N := \{0 \leq E_i \leq M\}$  such that  $\mu(A) > 1 - \epsilon/4$ . By the definition of  $A$ , clock rates for initial values in  $A$  are uniformly bounded. Therefore, one can find a sufficiently small  $\delta > 0$ , such that  $\mathbb{P}[\text{no clock rings on } [0, \delta]] \geq 1 - \epsilon/4$ . For any set  $U \subset \mathbb{R}_+^N$ , we have

$$\begin{aligned} (\mu P^\delta)(U) &= \int_{\mathbb{R}_+^N} P^\delta(\mathbf{E}, U) \mu(d\mathbf{E}) \\ &= \int_{A \cap U} P^\delta(\mathbf{E}, U) \mu(d\mathbf{E}) + \int_{A^c \cap U} P^\delta(\mathbf{E}, U) \mu(d\mathbf{E}) + \int_{A^c} P^\delta(\mathbf{E}, U) \mu(d\mathbf{E}) \\ &= \mu(A \cap U) - a_1 + a_2 + a_3, \end{aligned}$$

where

$$\begin{aligned} a_1 &= \int_{A \cap U} (1 - P^\delta(\mathbf{E}, U)) \mu(d\mathbf{E}) \leq \frac{\epsilon}{4} \mu(A \cap U) \leq \frac{\epsilon}{4} \\ a_2 &= \int_{U-A} P^\delta(\mathbf{E}, U) \mu(d\mathbf{E}) \leq \frac{\epsilon}{4} \mu(A - U) \leq \frac{\epsilon}{4} \\ a_3 &= \int_{A^c} \frac{\epsilon}{4} \mu(A \cap U) \leq \frac{\epsilon}{4} \mu(A^c) \leq \frac{\epsilon}{4}. \end{aligned}$$

In addition we have  $\mu(U) - \mu(A \cap U) \leq \mu(A^c) < \frac{\epsilon}{4}$ . Hence

$$|(\mu P^\delta)(U) - \mu(U)| < \epsilon$$

for any  $U \subset \mathbb{R}_+^N$ . By the definition of the total variation norm, we have

$$\|\mu P^\delta - \mu\| \leq \epsilon.$$

This completes the proof. □

We haven't talked about uniqueness so far. Usually the uniqueness of the invariant probability measure follows from the fact that  $P(x, \cdot)$  admits positive density everywhere.

**Proposition 5.7.** *For any  $h > 0$ ,  $\mathbf{E}_n^h$  admits at most one invariant probability measure.*

*Proof.* By the proof of Theorem 5.1, for any  $\mathbf{E} \in K$ ,  $P^{h/2}(\mathbf{E}, \cdot)$  has strictly positive density on  $K$ . In addition,  $P^{h/2}(\mathbf{E}_0, K) > 0$  for any  $\mathbf{E}_0 \in \mathbb{R}_+^N$ . Hence  $P^h(\mathbf{E}_0, \cdot)$  has positive density on  $K$ . This implies that every  $\mathbf{E}_0 \in \mathbb{R}_+^N$  belongs to the same ergodic component. Therefore  $\mathbf{E}_n^h$  cannot have more than one invariant probability measures. □

In summary, we have the following conclusions for  $\mathbf{E}_t$ .

- (1) For any  $T_L, T_R$ , there exists a unique invariant probability measure  $\pi$ , i.e., the nonequilibrium steady-state, which is absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}_+^N$ .
- (2) For almost every  $\mathbf{E}_0 \in \mathbb{R}_+^N$  and any sufficiently small  $\epsilon > 0$ , we have

$$\lim_{t \rightarrow \infty} t^{2-\epsilon} \|\delta_{\mathbf{E}_0} P^t - \pi\|_{TV} = 0.$$

- (3) For any functions  $\eta, \xi \in L^\infty(\mathbb{R}_+^N)$ , we have

$$C_\pi^{\eta, \xi}(t) \leq O(1) \cdot t^{\epsilon-2}$$

for any  $\epsilon > 0$ .

## 6. EXAMPLE: RANDOM HALVES MODEL

**6.1. Derivation from deterministic dynamics.** Another way to simplify the multi-body problem as described in Figure 1 at the beginning of the previous section is to assume that particles do not interact directly. Instead, we divide the tube into a chain of  $N$  cells, each of which contains a rotating disk that plays the role of the “local environment”. As seen in Figure 7, particles can only exchange energy with the rotating disk. Then we connect this chain with two thermalized ends, called heat baths, such that thermalized particles can be injected into the system and particles in the system can exit by entering the heat bath. This is the Hamiltonian model proposed in [8]. We refer [8, 17] for details.

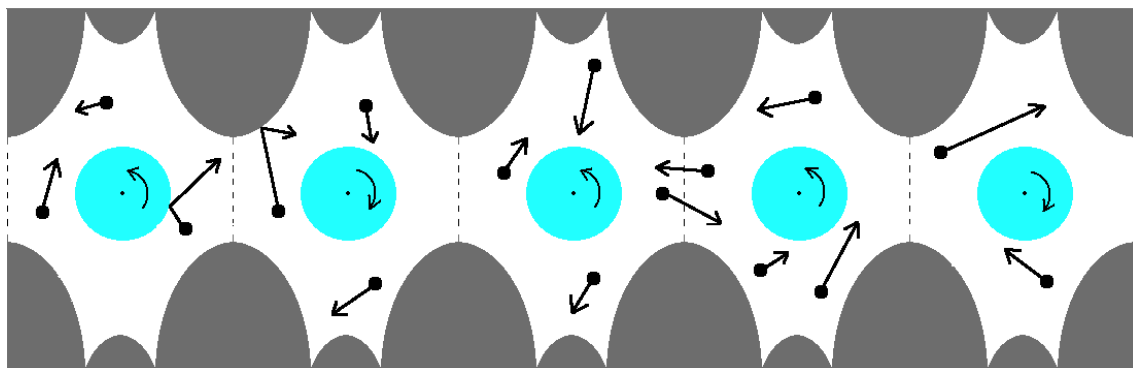


FIGURE 7. Moving particles do not interact with each other directly. There is a rotating disk with fixed center in each cell. The collision between particles and disks are determined by the conservation of kinetic energy and angular momentum.

A particle in this Hamiltonian model has chaotic trajectories and quick loss of “memory”. Therefore, it is natural to assume that the movement of each particle is stochastic, i.e., governed by an energy dependent exponential clock. When a clock associated with a particle rings, the particle either jumps to neighboring cells or exchanges energy with the local environment. The probability of occurrence of either event is a constant determined by the system. This reduces the Hamiltonian model to the so-called random halves model, which is described in the next subsection. We refer [17] for the full detail of this model reduction process and a brief justification of the model reduction.

**6.2. Model Description.** Consider  $N$  linearly ordered lattice sites  $\{1, 2, \dots, N\}$ , each containing an energy tank and storing a finite number of particles with certain amount of energy. The lattice sites are connected to two heat baths at the ends, denoted as sites 0 and  $N + 1$  for the sake of simplicity. The heat baths have temperatures  $T_L$  and  $T_R$ , as well as exponential particle injection rates of  $\rho_L$  and  $\rho_R$ , respectively. Particle energies are random variables with i.i.d. distributions with a probability density function

$$\frac{2}{\sqrt{\pi}T^{3/2}}\sqrt{x}e^{-\frac{x}{T}}$$

where  $T$  is the temperature of the heat bath from which the particle is emitted. Notice that when a particle is emitted by the left heat bath, it instantaneously appears at site 1. When a particle is emitted by the right heat bath, it instantaneously appears at site  $N$ .

An exponential clock is associated with each particle in the system, with rate  $(1+m)S\sqrt{x}$ , where  $x$  is the energy of the particle and  $m$  and  $S$  are system constants. When the clock of a particle rings, the particle jumps with probability  $\frac{1}{1+m}$  and “mixes” with probability  $\frac{m}{1+m}$ . Rules for jumping and “mixing” for a particle at site  $i$  carrying energy  $x_j^i$ , where  $j$  is the index of the particle in its site, are as follows.

When a particle jumps, it goes to either site  $i+1$  or site  $i-1$  with equal probability of  $\frac{1}{2}$ . Let  $k_i$  be the number of particles at site  $i$ . Then  $k_i$  decreases by 1 while the site that the particle jumps to has an increase in particle number by 1. Notice that the particle leaves the system if it jumps to the left or right heat bath. For “mixing”, we mean a particle exchanges energy with the stored energy at its site. Let  $s_i$  be the stored energy at site  $i$  and let  $x_j$  be the corresponding particle energy. The rule of energy exchange is  $(s'_i, x'_j) = (x_j u^2, s_i + x_j(1 - u^2))$ , where  $u$  is a uniform random variable distributed on  $(0, 1)$ .

Random halves model generates a Markov jump process

$$\omega_t = ((s_1(t), \{x_1^1(t), \dots, x_{k_1}^1(t)\}), \dots, (s_N(t), \{x_1^N(t), \dots, x_{k_N}^N(t)\})),$$

where  $k_1, \dots, k_N$  take values in  $\mathbb{N} \cup \{0, \infty\} = \{0, 1, \dots, \infty\}$ . The stochastic process takes values in the state space  $\Omega = \prod_{i=1}^N \Omega^i$ , where  $\Omega^i = \cup_{k=0}^{\infty} \Omega_k^i$ , and  $\Omega_k^i = \{(s_i, \{x_1^i, \dots, x_k^i\}) | s_i, x_j^i > 0\}$ . Since we regard particles as indistinguishable to avoid confusion when particles re-enter the system, we use unordered lists denoted by curly brackets. Notice that  $\Omega_k^i = \mathbb{R}_+ \times ((\mathbb{R}_+)^k / \sim)$ , where  $\sim$  is the equivalence relation given by  $\{x_1, \dots, x_k\} \sim \{x_{\sigma_1}, \dots, x_{\sigma_k}\}$ ,  $\sigma$  is any  $k$ -permutation.

Therefore, we can define the Markov jump process  $\omega(t)$  generated by random halves model on  $\Omega$ . We denote the transition kernel of  $\omega(t)$  by  $P^t(\omega, \cdot)$ . Let  $h > 0$  be a fixed number that represents the step size. The time- $h$  sampling chain of  $\omega_t$  is denoted by  $\omega_n^h$ , or simply  $\omega_n$  when it does not lead to a confusion. The transition kernel of  $\omega_n$  is denoted by  $P(\omega, \cdot)$ .

For the sake of later use, we will define a reference measure  $\Lambda$  on  $\Omega$ , where  $\Lambda = \prod_{i=1}^N \Lambda_i$  and  $\Lambda_i$  is the natural reference measure on  $\Omega^i$ , such that the restriction of  $\Lambda_i$  on  $\Omega_k^i$  is the quotient of the Lebesgue measure on  $\mathbb{R} \times \mathbb{R}^k$  under the relationship  $\sim$ .

**6.3. Verifying (A1) and (A2).** We will first work on the time- $h$  sampling chain of  $\omega_n$ . The verification of analytical conditions of  $\omega_n$  is based on the following theorem.

**Theorem 6.1.** *For any set  $K \subset \Omega$ , of the form  $\{\omega \in \Omega | 0 \leq k_i \leq K_0, 0 \leq s_i \leq S_0, c \leq x_j^i \leq C, i = 1 \sim N, j = 1 \sim k_i\}$ , where  $K_0, S_0, c, C$  are positive constants, and any  $h > 0$ , there exists a constant  $\eta > 0$  such that*

$$P(\omega, \cdot) > \eta \Lambda_K(\cdot),$$

for all  $\omega \in K$ , where  $\Lambda_K$  is the reference measure restricted to  $K$ .



*Proof.* For  $\omega = ((\hat{s}_1, \{\hat{x}_1^1, \dots, \hat{x}_{k_1}^1\}), \dots, (\hat{s}_N, \{\hat{x}_1^N, \dots, \hat{x}_{k_N}^N\})) \in K$ , let  $A_\omega(ds) = \{(s_1, \{x_1^1, \dots, x_{k_1}^1\}), \dots, (s_N, \{x_1^N, \dots, x_{k_N}^N\}) \mid \hat{s}_1 \leq s_1 \leq \hat{s}_1 + ds, \dots, \hat{s}_N \leq s_N \leq \hat{s}_N + ds, \hat{x}_j^i \leq x_j^i \leq \hat{x}_j^i + ds, \text{ for } 1 \leq i \leq N, 1 \leq j \leq k_i\}$  for  $ds \ll 1$ .

It then suffices to prove that for each

$$\bar{\omega} = ((\bar{s}_1, \{\bar{x}_1^1, \dots, \bar{x}_{k_1}^1\}), \dots, (\bar{s}_N, \{\bar{x}_1^N, \dots, \bar{x}_{k_N}^N\})) \in K,$$

we have

$$P(\bar{\omega}, A_\omega(ds)) \geq \sigma ds^{1+k_1} \dots ds^{1+k_N}$$

for  $0 < ds \ll 1$ , where  $\sigma$  is a strictly positive constant that is independent of  $\bar{\omega}$  and  $\omega$ .

In order to do so, we will construct a sequence of events to go from the state  $\bar{\omega}$  to  $A_\omega(ds)$  with positive probability. Let  $M = \sum_{i=1}^N k_i$  and  $\bar{M} = \sum_{i=1}^N \bar{k}_i$ . Let  $\delta = \frac{h}{M+\bar{M}+N}$ . And let  $E = \sum_{i=1}^N (\hat{s}_i + ds)$ . We consider the events  $F_i, G, H_i, A_i$ , where  $i = 1 \sim N$ .

- $F_i = \{\text{on } ((i-1)\delta, i\delta], i = 1 \sim \bar{M}, \text{ no new particle enters, and one particle present initially exits the system without exchanging energy}\}$ .
- Define an auxiliary event  $A_i = \{\hat{s}_i \leq s_i \leq \hat{s}_i + ds\}$  on  $((M+i-1)\delta, (M+i)\delta]$ .
- $H_i = A_i \cap \{\text{on } (M\delta + (i-1)\delta, M\delta + i\delta], \text{ exactly one particle, carrying energy on the interval } [E, 2E], \text{ enters from the left and jumps through all sites } s_j, j < i \text{ without exchanging energy, then exchanges energy with site } i, \text{ and jumps to exit the system from the right}\} \cap \{\text{no other new particle enters}\}$  for  $i = M \sim M+N$ .
- $G = \{\text{on } (i\delta, (i+1)\delta], i = \bar{M} + N \sim \bar{M} + N + M, k_i \text{ particles with energy } x_j^i \in [\hat{x}_j^i, \hat{x}_j^i + ds] \text{ for } j = 1 \sim k_i \text{ enters the system from the left and jumps until reaching site } i, \text{ without exchanging energy}\} \cap \{\text{no new particle enters and existing particles do nothing}\}$ .

The idea is that particles initially present at each site are first emptied from the system. Then for each site, one particle with sufficiently large amount of energy enters the system to mix at the corresponding site. Lastly, particles in the target set enter the system and jump to corresponding sites. We need to show that the probability of occurrence of the sequence of events described above is always strictly positive. Here are the considerations.

- (1) The initial number of particles  $k_i \leq K_0$  for each  $i = 1 \sim N$ . Clock rates  $(1+m)S\sqrt{x_j^i}$  are bounded above zero since  $x_j^i \geq c$ , and bounded below infinity since  $x_j^i \leq C$  by assumption. Hence  $F_i$  occurs with strictly positive probability.
- (2) Let  $u \in (0, 1)$  be the fraction in the mixing that puts  $s_i \in [\hat{s}_i, \hat{s}_i + ds]$ . Let  $X$  be the particle energy in the event  $H_i$ . Rearrange the terms to get  $\hat{s}_i/X \leq u^2 \leq (\hat{s}_i + ds)/X$ . Note that  $X$  is bounded from above by  $2NS_0$ . Therefore,  $\mathbb{P}[A_i] > \alpha ds$  for some constant  $\alpha > 0$ .
- (3) In addition, during the event  $\{H_i, i = 1 \sim N\}$ ,  $E$  is bounded from above by  $NS_0$ . The probability of  $H_i$  is greater than  $\text{const} \cdot ds$ .

- (4) The number of particles in the destination set  $\bar{k}_i \leq K_0$  for each  $i = 1 \sim N$ , and clock rates are bounded both above from zero and below from  $C$ . Hence  $G$  occurs with probability at least  $\text{const} \cdot ds^{k_1} \dots ds^{k_N}$ .

In addition, all these probabilities and probability densities are uniformly bounded from below for all  $\bar{\omega}$  and  $\omega$  in  $K$ .  $\square$

As a corollary, we can prove that  $\omega_n$  is both aperiodic and irreducible with respect to the reference measure.

**Corollary 6.2.**  $\omega_n$  is a strongly aperiodic Markov chain.

*Proof.* By 6.1,  $K$  is a uniform reference set. In addition  $\Lambda_K(K) > 0$ . The strong aperiodicity follows from its definition.  $\square$

Therefore  $\omega_n$  is aperiodic.

**Corollary 6.3.**  $\omega_n$  is  $\Lambda$ -irreducible, where  $\Lambda$  is the reference measure on  $\Omega$ .

*Proof.* Let  $A \subset \Omega$  be a set with strictly positive measure. Then there exists a set  $U$  that has the form  $U = \{\omega \in \Omega | 0 \leq k_i \leq K_0, 0 \leq s_i \leq S_0, c \leq x_j^i \leq C, i = 1 \sim N, j = 1 \sim k_i\}$  and  $\Lambda_K(U \cap A) > 0$ .

For any  $\omega_0 \in \Omega$  and the time step  $h > 0$ , there exists a set  $K$  that has the form as in Theorem 6.1 such that  $\omega_0 \in K$ . Hence same construction as in Theorem 6.1 implies that  $P^h(\omega_0, \cdot) > \eta \Lambda_K(\cdot)$  for some  $\eta > 0$ . Therefore  $P^h(\omega_0, A) > \eta \Lambda_K(A) > 0$ .  $\square$

**6.4. Absolute continuity of invariant measure.** The proof of the absolute continuity of  $\pi$  with respect to  $\Lambda$  is similar as in the previous section.

**Proposition 6.4.** If  $\pi$  is an invariant measure of  $\omega_t$ , then  $\pi$  is absolutely continuous with respect to  $\lambda$  with a strictly positive density.

*Proof.* Let  $\pi = \pi_{abs} + \pi_{\perp}$  be an invariant measure, where  $\pi_{abs}$  and  $\pi_{\perp}$  are absolutely continuous and singular components with respect to  $\Lambda$  respectively. Since  $\pi_{abs} \ll \Lambda$ ,  $\pi_{abs} P^t \ll \Lambda$  for any  $t > 0$  by Lemma 6.3 of [17]. For similar reasons as in the previous model, we again refer readers to [17] for detailed calculations. The rest of the proof then follows the same line as in the proof of Proposition 5.4.  $\square$

**6.5. Verifying (N1) and (N2).** Now we will present our numerical results for the random halves model. We let  $N = 3$  or  $4$ , depending on the computational cost of the simulation. The uniform reference set  $\mathfrak{C}$  is chosen as

$$\mathfrak{C} = \{\omega \in \Omega | 0 \leq k_i \leq 40, 0 \leq s_i \leq 100, 0.1 \leq x_j^i \leq 100, i = 1 \sim N, j = 1 \sim k_i\}$$

Throughout our numerical justification about  $\tau_{\mathfrak{C}} = \tau_{\mathfrak{C}}(h)$ , we let  $h = 0.1$ . We verify (N1) for the numerically generated invariant measure. The numerical invariant measure is generated by running the process for a sufficiently long time from a suitable initial distribution. In our simulation, the initial distribution for the simulation of the numerical invariant measure is

$$\omega_0 \sim ((s_1(t), \{x_1^1(t), \dots, x_{k_1}^1(t)\}), \dots, (s_N(t), \{x_1^N(t), \dots, x_{k_N}^N(t)\})),$$

where  $N = 3$ , each  $s_i$  is uniformly distributed between 0 and 100, each  $k_i$  is a poisson distribution with mean  $(\rho_L + \rho_R)/2$ , and each  $x_j^i$  is an exponential distribution with mean  $(T_L + T_R)/2$ . We find that when  $T = 100$ , the expectation of the observables we have tested are stabilized (Figure 8). Therefore the numerical invariant measure is chosen as  $\mu_0 P^{100}$ . Then from the result of our simulation that chooses the numerical invariant measure as the specific starting state, we conclude that  $\mathbb{P}_{\omega_0}[\tau_{\mathfrak{C}} > t] \sim t^{-2}$  and  $\mathbb{P}_{\pi}[\tau_{\mathfrak{C}} > t] \sim t^{-2}$ . (Figure 9).

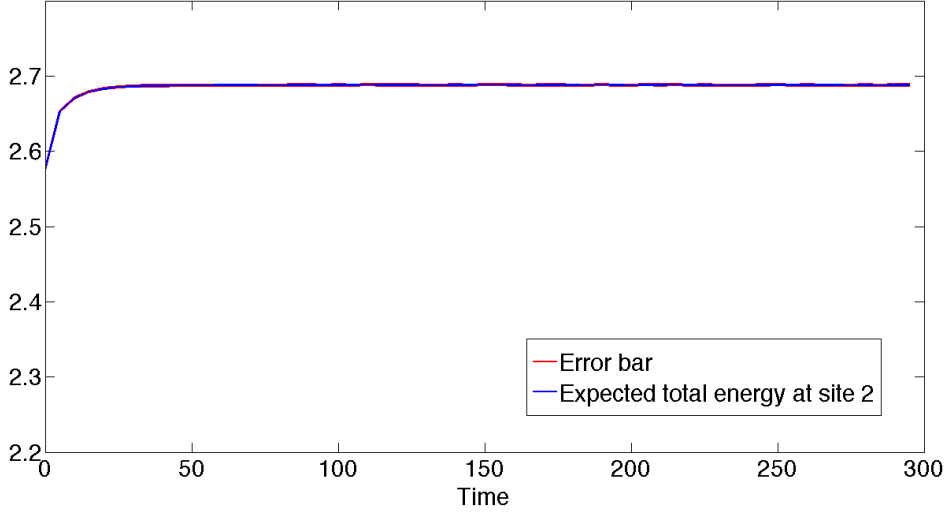


FIGURE 8. Total energy at the second site verses time. Sample size  $= 1 \times 10^8$ .

In order to verify **(N2)**, we look for the maximal of  $f(x) := \mathbb{E}_{\omega}[\tau_{\mathfrak{C}}^{2-\epsilon}]$ . Figure 10 shows that when  $\epsilon = 0.3$ ,  $f(x)$  increases monotonically with number of particles at each site, and decreases monotonically with decreasing site energy. Therefore we believe the maximal of  $f(x)$  in  $\mathfrak{C}$  is  $\omega_* = \{(0, \{0.1, \dots, 0.1\}), (0, \{0.1, \dots, 0.1\}), (0, \{0.1, \dots, 0.1\})\}$ , for  $k_i = 40, i = 1 \sim 3$ . For the same reason as in Section 5, we did not use numerical optimization tools to find  $\omega_*$ .

**6.6. Main Conclusion.** The previous subsection verifies **(N1)** for  $\delta_{\omega_0}$  and  $\pi$ , as well as **(N2)**. The slopes of  $\mathbb{P}[\tau_{\mathfrak{C}} > t]$  for both initial conditions are 2. Note that  $\mathbb{P}[\Psi_t = \Psi_0 | \Psi_0 \in K] = \mathbb{P}_{\Psi_0}[\text{no clock rings up to } t]$  is uniformly positive for each given  $t > 0$ . By Proposition 4.5, **(N1)** and **(N2)** hold for  $\omega_n$  with parameter  $2 - \epsilon/2$  for arbitrarily small  $\epsilon > 0$ . Therefore, conclusions (a)-(d) in Section 4.4 holds for  $\omega_n$ .

It remains to pass the results for  $\omega_n$  to  $\omega_t$ . By Proposition 4.3, it is sufficient to prove “continuity at zero” for  $\omega_t$ .

**Lemma 6.5.** *For any probability measure  $\mu$  on  $\Omega$ ,*

$$\lim_{\delta \rightarrow 0} \|\mu P^{\delta} - \mu\|_{TV} = 0$$

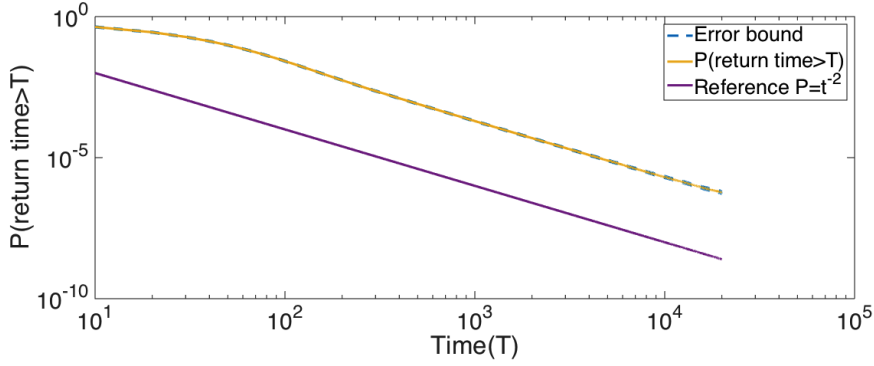


FIGURE 9.  $\mathbb{P}_\pi[\tau_{\mathfrak{C}} > t]$ ,  $\pi$  is the numerical steady state obtained from long time averaging. Sample size =  $1 \times 10^8$ . The purple line is a reference line with slope  $-2$ .

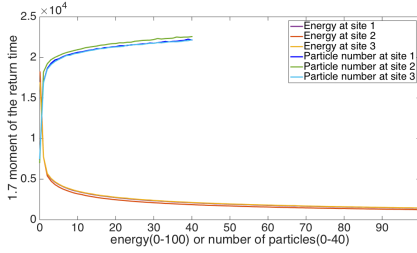


FIGURE 10. Change of  $\mathbb{E}_\omega[\tau_{\mathfrak{C}}^{1.7}]$  for varying  $\omega$  when particle energy and particle number change. Sample size =  $1 \times 10^8$  for each initial condition.

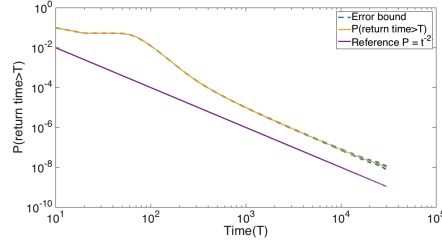


FIGURE 11.  $\mathbb{P}_{\omega_*}[\tau_{\mathfrak{C}} > t]$  for  $\omega_* = \{(0, \{0.1, \dots, 0.1\}), \dots, (0, \{0.1, \dots, 0.1\})\}$ , for  $k_i = 40, i = 1 \sim 3$ . Sample size =  $1 \times 10^{10}$ . The purple line is a reference line with slope  $-2$ .

*Proof.* It is sufficient to prove that for any  $\varepsilon > 0$ , there exists a  $\delta > 0$  such that

$$\|\mu P^\delta - \mu\|_{TV} \leq \varepsilon.$$

Since  $\mu$  is finite, there exists a bounded set  $A \subset \Omega := \{0 \leq k_i \leq K, 0 \leq s_i \leq S, 0 \leq x_j^i \leq M\}$  such that  $\mu(A) > 1 - \varepsilon/4$ . By the definition of  $A$ , clock rates for initial values in  $A$  are uniformly bounded. Therefore, one can find a sufficiently small  $\delta > 0$ , such that  $\mathbb{P}[\text{no clock rings on } [0, \delta]] \geq 1 - \varepsilon/4$ . For any set  $U \subset \Omega$ , the same calculation as in the proof of Lemma 5.6 implies

$$|(\mu P^\delta)(U) - \mu(U)| < \varepsilon$$

for any  $U \subset \Omega$ . By the definition of the total variation norm, we have

$$\|\mu P^\delta - \mu\| \leq \varepsilon.$$

This completes the proof.  $\square$

It remains to prove the uniqueness of the invariant measure.

**Proposition 6.6.** *For any  $h > 0$ ,  $\omega_n^h$  admits at most one invariant probability measure.*

*Proof.* This proof is the same as that of Proposition 5.7.  $\square$

In summary, we have the following conclusions for  $\omega_t$ .

- (1) For any  $T_L, T_R$ , there exists a unique invariant probability measure  $\pi$ , i.e., the nonequilibrium steady-state, which is absolutely continuous with respect to the reference measure on  $\Omega$ .
- (2) For almost every  $\omega_0 \in \Omega$  and any sufficiently small  $\varepsilon > 0$ , we have

$$\lim_{t \rightarrow \infty} t^{2-\varepsilon} \|\delta_{\omega_0} P^n - \pi\|_{TV} = 0.$$

- (3) For any functions  $\eta, \xi \in L^\infty(\Omega)$ , we have

$$C_\pi^{\eta, \xi}(t) \leq O(1) \cdot t^{\varepsilon-2}$$

for any  $\varepsilon > 0$ .

## REFERENCES

- [1] Federico Bonetto, Joel L Lebowitz, and Luc Rey-Bellet. Fourier's law: a challenge to theorists. *Mathematical physics*, 2000:128–150, 2000.
- [2] Jean Bricmont and Antti Kupiainen. Towards a derivation of Fourier's law for coupled anharmonic oscillators. *Communications in mathematical physics*, 274(3):555–626, 2007.
- [3] Leonid Bunimovich, Carlangelo Liverani, Alessandro Pellegrinotti, and Yurii Suhov. Ergodic systems of  $n$  balls in a billiard table. *Communications in mathematical physics*, 146(2):357–396, 1992.
- [4] Noé Cuneo and J-P Eckmann. Non-equilibrium steady states for chains of four rotors. *Communications in Mathematical Physics*, pages 1–37, 2016.
- [5] Noé Cuneo, Jean-Pierre Eckmann, and Christophe Poquet. Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors. *Nonlinearity*, 28(7):2397, 2015.
- [6] B Derrida, JL Lebowitz, and ER Speer. Large deviation of the density profile in the steady state of the open symmetric simple exclusion process. *Journal of statistical physics*, 107(3-4):599–634, 2002.
- [7] J-P Eckmann, C-A Pillet, and Luc Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Communications in Mathematical Physics*, 201(3):657–697, 1999.
- [8] J-P Eckmann and L-S Young. Nonequilibrium energy profiles for a class of 1-d models. *Communications in Mathematical Physics*, 262(1):237–267, 2006.
- [9] Pierre Gaspard and Thomas Gilbert. Heat conduction and Fourier's law by consecutive local mixing and thermalization. *Physical review letters*, 101(2):020601, 2008.
- [10] Alexander Grigo, Konstantin Khanin, and Domokos Szasz. Mixing rates of particle systems with energy exchange. *Nonlinearity*, 25(8):2349, 2012.
- [11] Martin Hairer. Convergence of Markov processes. *lecture notes*, 2010.
- [12] Martin Hairer and Jonathan C Mattingly. Slow energy dissipation in anharmonic oscillator chains. *Communications on Pure and Applied Mathematics*, 62(8):999–1032, 2009.
- [13] Søren F Jarner, Gareth O Roberts, et al. Polynomial convergence rates of Markov chains. *The Annals of Applied Probability*, 12(1):224–247, 2002.
- [14] C Kipnis, C Marchioro, and E Presutti. Heat flow in an exactly solvable model. *Journal of Statistical Physics*, 27(1):65–74, 1982.

- [15] Yao Li. On the stochastic behaviors of locally confined particle systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(7):073121, 2015.
- [16] Yao Li. On the polynomial convergence rate to nonequilibrium steady-states. *arXiv preprint arXiv:1607.08492*, 2016.
- [17] Yao Li and Lai-Sang Young. Nonequilibrium steady states for a class of particle systems. *Nonlinearity*, 27(3):607, 2014.
- [18] Yao Li and Lai-Sang Young. Polynomial convergence to equilibrium for a system of interacting particles. *Annals of Applied Probability*, accepted.
- [19] Torngny Lindvall. *Lectures on the coupling method*. Courier Dover Publications, 2002.
- [20] Kerrie L Mengersen, Richard L Tweedie, et al. Rates of convergence of the hastings and metropolis algorithms. *The Annals of Statistics*, 24(1):101–121, 1996.
- [21] Sean P Meyn and Richard L Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, 2009.
- [22] Esa Nummelin. A splitting technique for harris recurrent markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 43(4):309–318, 1978.
- [23] Esa Nummelin and Pekka Tuominen. The rate of convergence in orey’s theorem for harris recurrent markov chains with applications to renewal theory. *Stochastic Processes and Their Applications*, 15(3):295–311, 1983.
- [24] Luc Rey-Bellet and Lawrence E Thomas. Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Communications in Mathematical Physics*, 215(1):1–24, 2000.
- [25] Pekka Tuominen and Richard L Tweedie. Subgeometric rates of convergence of f-ergodic markov chains. *Advances in Applied Probability*, pages 775–798, 1994.
- [26] Tatiana Yarmola. Sub-exponential mixing of open systems with particle–disk interactions. *Journal of Statistical Physics*, pages 1–20, 2013.
- [27] Tatiana Yarmola. Sub-exponential mixing of random billiards driven by thermostats. *Nonlinearity*, 26(7):1825, 2013.

YAO LI: DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MASSACHUSETTS  
AMHERST, AMHERST, MA, 01002, USA

*E-mail address*: yaoli@math.umass.edu

HUI XU: DEPARTMENT OF MATHEMATICS, AMHERST COLLEGE, AMHERST, MA, 01002,  
USA

*E-mail address*: huxu18@amherst.edu